



Department of Digital Business

**Journal of Artificial Intelligence and Digital Business (RIGGS)**

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 2 (2026) pp: 15540-15560

P-ISSN: 2963-9298, e-ISSN: 2963-914X

---

## Analisis Sentimen Masyarakat terhadap Kebijakan Gubernur Jawa Barat di Sosial Media menggunakan *Algoritma Long Short Term Memory*

M Fikri Setiawan, Elvi Rahmi

Teknik informatika, Politeknik negeri Bengkalis

[setiawanmfikri@gmail.com](mailto:setiawanmfikri@gmail.com), [elviraahmi@polbeng.ac.id](mailto:elviraahmi@polbeng.ac.id)

### Abstrak

Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap kebijakan Gubernur Jawa Barat yang terekam pada berbagai media sosial, seperti Twitter, Instagram, dan TikTok, dengan memanfaatkan algoritma Long Short-Term Memory (LSTM). Analisis sentimen dipilih karena mampu memberikan gambaran mengenai persepsi publik secara cepat dan berkelanjutan seiring meningkatnya penggunaan media sosial sebagai sarana penyampaian opini. Permasalahan utama penelitian ini adalah karakteristik opini masyarakat yang sangat dinamis, berjumlah besar, serta tersebar di berbagai platform digital, sehingga metode konvensional seperti survei atau wawancara memiliki keterbatasan dalam memperoleh informasi secara real time dan representatif. Dataset penelitian diperoleh melalui proses scrapping komentar berbahasa Indonesia yang berkaitan dengan kebijakan Gubernur Jawa Barat. Data kemudian diproses menggunakan tahapan Knowledge Discovery in Databases (KDD) yang meliputi seleksi data, preprocessing, transformasi data, pelabelan sentimen, serta pembobotan kata menggunakan metode Word Embedding untuk menghasilkan representasi fitur yang optimal. Selanjutnya, model LSTM dilatih untuk mengklasifikasikan sentimen ke dalam tiga kategori, yaitu positif, negatif, dan netral. Evaluasi performa model dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score guna mengukur kemampuan klasifikasi secara menyeluruh. Hasil penelitian menunjukkan bahwa algoritma LSTM mampu mengidentifikasi dan mengklasifikasikan sentimen masyarakat dengan tingkat performa yang baik, sehingga dapat menggambarkan kecenderungan opini publik terhadap berbagai kebijakan yang diterapkan. Temuan ini diharapkan dapat menjadi landasan dalam pengembangan sistem pemantauan opini publik berbasis kecerdasan buatan yang mendukung proses evaluasi kebijakan serta pengambilan keputusan pemerintah daerah secara lebih cepat, responsif, objektif, dan tepat sasaran.

*Kata kunci:* Analisis Sentimen, LSTM, Kebijakan Gubernur Jawa Barat, Media Sosial

### 1. Latar Belakang

Selama ini, evaluasi penerimaan kebijakan publik di Jawa Barat umumnya dilakukan melalui metode survei tatap muka atau wawancara langsung. Misalnya, pedoman survei yang berlaku menyebutkan bahwa salah satu teknik utama pengumpulan data adalah kuesioner dengan wawancara tatap muka [1]. Selain itu, survei Indikator Politik Indonesia terhadap tingkat kepuasan publik di Jawa Barat dilaksanakan menggunakan metode wawancara tatap muka, melibatkan 600 responden pada Mei 2025 [2]. Meskipun metode ini mampu menggali informasi secara mendalam, prosesnya membutuhkan waktu, biaya, dan sumber daya yang besar, serta rentan terhadap bias responden.

Di era digital, opini dan sentimen publik terhadap suatu kebijakan dapat berubah secara signifikan hanya dalam hitungan jam, dipengaruhi oleh arus informasi dan interaksi di media sosial [3]. Tanpa adanya pemantauan sentimen secara real-time, pemerintah daerah berisiko terlambat mengidentifikasi potensi penolakan kebijakan maupun peluang untuk meningkatkan dukungan publik [4]. Penelitian ini menawarkan solusi berbasis *deep learning* yang mampu mengolah ribuan opini masyarakat dari media sosial secara otomatis dan akurat, dengan mengklasifikasikan sentimen ke dalam tiga kategori: positif, negatif, dan netral. Pendekatan ini berpotensi dimanfaatkan sebagai *early warning system* bagi pemerintah daerah untuk mendeteksi dinamika opini publik dan merumuskan strategi respons kebijakan yang tepat waktu [5], [6].

Di Indonesia, dengan penetrasi internet yang mencapai 77% dari total populasi [7] dan lebih dari 167 juta pengguna media sosial aktif [8], platform-platform seperti Twitter, Instagram, dan TikTok menyediakan representasi signifikan dari opini masyarakat. Aktivitas pengguna di ketiga platform ini secara konsisten

mencerminkan dinamika sosial dan politik yang berkembang, termasuk reaksi terhadap kebijakan pemerintah baik di tingkat nasional maupun daerah.

Provinsi Jawa Barat dipilih sebagai fokus penelitian karena jumlah penduduknya yang terbesar di Indonesia, serta perannya yang penting di berbagai bidang seperti ekonomi, sosial, politik, dan digital. Dalam aspek ekonomi, Jawa Barat berkontribusi sekitar 23,4% terhadap Produk Domestik Bruto (PDB) nasional. Oleh karena itu, kebijakan yang diterapkan di sini bisa berdampak besar terhadap perekonomian Indonesia secara keseluruhan [10]. Dari segi politik, Jawa Barat dikenal sebagai salah satu wilayah dengan suara terbanyak dalam pemilu. Kecenderungan pendapat publik di sini sering menjadi petunjuk arah politik nasional [12]. Di sisi digital, Jawa Barat memiliki tingkat penggunaan internet dan media sosial yang tinggi, menjadikannya salah satu wilayah dengan aktivitas online paling banyak di Indonesia [13]. Kondisi ini menciptakan peluang sekaligus tantangan dalam memahami perubahan opini publik yang terjadi begitu cepat di media sosial. Selain itu, keberagaman etnis dan budaya masyarakat Jawa Barat memengaruhi keragaman pandangan dan respons masyarakat terhadap kebijakan pemerintah setempat [11].

Kebijakan-kebijakan yang dikeluarkan oleh Pemerintah Provinsi Jawa Barat mencakup spektrum luas seperti infrastruktur, ekonomi, sosial, pendidikan, hingga lingkungan. Setiap kebijakan tersebut memicu respons masyarakat yang terekam dalam bentuk komentar dan unggahan di Twitter, Instagram, dan TikTok, menghasilkan *big data* sosial yang kaya akan informasi terkait aspirasi, kritik, maupun dukungan publik. Menganalisis data ini dapat memberikan *insight* yang berharga mengenai persepsi masyarakat secara real-time dan berbasis bukti untuk mendukung pengambilan keputusan kebijakan yang lebih responsif [12].

Namun, metode analisis sentimen tradisional seperti survei manual memiliki keterbatasan dalam hal kecepatan, skala, konsistensi, dan potensi bias. Oleh karena itu, pendekatan berbasis kecerdasan buatan khususnya *Deep Learning* menawarkan solusi yang lebih adaptif dan skalabel. Salah satu algoritma yang menonjol adalah Long Short-Term Memory (LSTM), yang dirancang khusus untuk memproses data sekuensial seperti teks dan mempertahankan konteks jangka panjang. Dengan kemampuan ini, LSTM sangat relevan untuk menangani teks dalam bahasa Indonesia yang memiliki struktur morfologi dan sintaksis kompleks, termasuk penggunaan bahasa campuran dan *code-switching* seperti dalam tweet masyarakat Jawa Barat.

Sebagian besar penelitian analisis sentimen di Indonesia masih berfokus pada satu platform media sosial saja atau menggunakan metode pembelajaran mesin konvensional seperti Naïve Bayes dan Support Vector Machine (SVM) [13], [14]. Pendekatan tersebut sering kali kurang optimal dalam menangkap konteks linguistik yang kompleks, sehingga akurasi terbatas. Selain itu, penelitian terdahulu umumnya tidak secara spesifik mengkaji sentimen masyarakat terhadap kebijakan di tingkat provinsi yang memiliki kompleksitas demografis seperti Jawa Barat. Penelitian ini mengisi kekosongan tersebut dengan menggabungkan data dari tiga platform media sosial utama (Twitter, Instagram, dan TikTok) serta menerapkan algoritma LSTM untuk memperoleh analisis sentimen yang lebih akurat, komprehensif, dan relevan dengan konteks lokal.

Penelitian ini diharapkan dapat menjadi acuan pengembangan sistem pemantauan opini publik berbasis kecerdasan buatan yang mampu memberikan informasi strategis kepada pemerintah daerah, khususnya Pemerintah Provinsi Jawa Barat, untuk merumuskan kebijakan yang lebih responsif, tepat sasaran, dan berorientasi pada kebutuhan masyarakat.

## 2. Metode Penelitian

Penelitian ini menggunakan metode **Knowledge Discovery in Databases (KDD)** untuk menganalisis sentimen masyarakat terhadap kebijakan Gubernur Jawa Barat berdasarkan komentar pada media sosial Twitter, Instagram, dan TikTok. Data penelitian diperoleh melalui teknik *scraping* menggunakan **Instant Data Scraper** dengan kata kunci yang berkaitan dengan berbagai kebijakan Gubernur Jawa Barat. Dataset penelitian berjumlah **10.794 komentar** berbahasa Indonesia yang dikumpulkan dari ketiga platform tersebut.

Tahapan penelitian diawali dengan **data selection**, yaitu memilih komentar yang relevan dengan kebijakan pemerintah daerah, meliputi kebijakan pendidikan, sosial, ekonomi, dan lingkungan. Selanjutnya dilakukan **preprocessing** untuk meningkatkan kualitas data melalui tahapan *case folding*, *cleaning*, *tokenization*, *stopwords filtering*, dan *stemming* sehingga diperoleh data teks yang bersih dan siap diolah.

Tahap berikutnya adalah **pelabelan sentimen** menggunakan pendekatan *hybrid*. Sebanyak **2.008 komentar** diberi label secara manual oleh guru Bahasa Indonesia melalui teknik *stratified random sampling*, sedangkan **8.796 komentar** lainnya diberi label secara otomatis menggunakan metode **lexicon-based** dengan kamus sentimen bahasa Indonesia yang telah disesuaikan dengan konteks kebijakan pemerintahan dan karakteristik bahasa media sosial. Sentimen dikelompokkan ke dalam tiga kategori, yaitu positif, negatif, dan netral.

Data yang telah diberi label kemudian memasuki tahap **transformation**, yaitu mengubah teks menjadi representasi numerik menggunakan **Word2Vec** berdimensi 128 sebagai *word embedding*. Dataset selanjutnya dibagi menjadi **80% data latih** dan **20% data uji**. Model klasifikasi dibangun menggunakan algoritma **Long Short-Term Memory (LSTM)** untuk mempelajari pola hubungan antar kata serta mengklasifikasikan sentimen komentar masyarakat.

Tahap akhir adalah **evaluation**, yaitu mengukur performa model menggunakan **confusion matrix**, **accuracy**, **precision**, **recall**, dan **F1-score**. Selain itu, dilakukan visualisasi distribusi sentimen melalui diagram batang untuk mengetahui kecenderungan opini masyarakat terhadap kebijakan Gubernur Jawa Barat. Seluruh proses pengolahan dan pelatihan model dilakukan menggunakan bahasa pemrograman Python pada lingkungan **Google Colab** dan **Jupyter Notebook**, sedangkan penyusunan laporan memanfaatkan aplikasi Microsoft Office.

### 3. Hasil dan Diskusi

#### a. Gambaran Umum Penelitian

Media sosial menjadi sarana utama masyarakat dalam menyampaikan opini terhadap berbagai kebijakan pemerintah, termasuk kebijakan Gubernur Jawa Barat. Respons yang muncul berupa dukungan, kritik, maupun pendapat netral yang tersebar melalui komentar dan diskusi pada berbagai platform digital. Penelitian ini bertujuan menganalisis sentimen masyarakat terhadap kebijakan Gubernur Jawa Barat menggunakan metode Long Short-Term Memory (LSTM), yaitu model *deep learning* yang mampu memahami pola dan konteks dalam rangkaian teks. Model tersebut mengklasifikasikan komentar ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral melalui tahapan pengumpulan data, prapemrosesan, pelabelan, hingga pengujian model. Hasil analisis diharapkan mampu memberikan gambaran mengenai persepsi masyarakat terhadap kebijakan yang diterapkan.

Data penelitian diperoleh melalui proses penambangan data (*data mining*) pada media sosial Instagram, Twitter (X), dan TikTok menggunakan ekstensi Instant Data Scraper. Alat ini digunakan untuk mengekstraksi komentar atau cuitan yang berkaitan dengan kebijakan Gubernur Jawa Barat berdasarkan kata kunci yang telah ditentukan, meliputi nama gubernur, nama kebijakan, serta istilah lain yang relevan. Data yang berhasil dikumpulkan kemudian disimpan dalam format CSV atau Excel agar mudah diproses menggunakan Python, selanjutnya digabungkan menjadi satu dataset sehingga dapat merepresentasikan opini masyarakat secara lebih luas.

#### b. Penegembangan Data

Tahapan penambangan data meliputi penentuan sumber data dan kata kunci, pengambilan data menggunakan Instant Data Scraper, penyimpanan hasil ekstraksi, serta penggabungan data dari seluruh platform media sosial. Dataset yang diperoleh selanjutnya digunakan pada tahap prapemrosesan dan analisis sentimen menggunakan metode LSTM. Contoh hasil penambangan data disajikan pada Tabel 1, yang dipilih berdasarkan relevansi terhadap kebijakan Gubernur Jawa Barat serta mewakili variasi sentimen positif, negatif, dan netral. Contoh tersebut juga menunjukkan bahwa beberapa komentar memiliki makna yang ambigu sehingga memerlukan analisis yang lebih mendalam menggunakan model LSTM.

Proses pengumpulan data dilakukan pada periode 20 Juli 2023 hingga 17 Agustus 2023. Dari hasil penambangan diperoleh 10.809 komentar, terdiri atas 6.142 komentar dari TikTok, 4.054 komentar dari Instagram, dan 613 komentar dari Twitter (X). Seluruh komentar tersebut menjadi dataset awal penelitian sebelum melalui proses prapemrosesan untuk menghapus data duplikat, komentar yang tidak relevan, serta akun dengan keterbatasan akses. Tahapan ini bertujuan menghasilkan dataset yang bersih dan representatif sehingga mampu mencerminkan opini publik terhadap kebijakan Gubernur Jawa Barat secara lebih akurat.

### c. Data Preprocessing

Selanjutnya setelah data didapatkan melakukan pemrosesan data melalui lima proses data diantaranya *cleansing*, *normalization*, *stopword removal*, *stemming*, dan *tokenization*.

#### 1) *Cleansing*

Proses *cleansing* data dilakukan untuk membersihkan teks komentar sebelum digunakan pada tahap analisis sentimen. Tahapan ini bertujuan untuk mengurangi *noise* pada data dengan cara menyeragamkan bentuk teks dan menghilangkan karakter yang tidak relevan. Pada proses ini, teks komentar diubah menjadi huruf kecil serta dibersihkan dari tautan, karakter khusus, angka, dan spasi berlebih.

Adapun tahapan *cleansing* data yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Mengimpor pustaka *pandas* untuk membaca dan mengelola data komentar dalam bentuk *DataFrame*, serta pustaka *regular expressions* (re) untuk melakukan pencarian dan penghapusan pola teks tertentu yang tidak diperlukan.
2. Menjalankan fungsi *case folding* untuk mengubah seluruh teks komentar menjadi huruf kecil (*lowercase*), menghapus tautan (*URL*), menghilangkan karakter non-alfabet, serta membersihkan spasi berlebih yang muncul akibat proses penghapusan karakter.
3. Menerapkan fungsi *case folding* pada kolom “komentar” dan menyimpan hasilnya ke dalam kolom baru bernama “clean\_text”. Selanjutnya, hasil *cleansing* data disimpan ke dalam berkas “hasil\_cleansing\_komentar.csv” agar dapat digunakan pada tahap prapemrosesan data selanjutnya.

```
def cleansing(text):  
    # mengubah seluruh huruf menjadi huruf kecil  
    # sehingga agar tidak ada perbedaan antara huruf besar dan huruf kecil  
    text = text.lower()  
  
    # menghapus tautan (url) yang dimulai dengan http atau https  
    # karena tautan tidak memiliki makna sentimen  
    text = re.sub('https?://.*', '', text)  
  
    # menghapus seluruh karakter selain huruf alfabet (a-z) dan spasi  
    # angka, tanda baca, dan simbol dianggap sebagai noise  
    text = re.sub('[^a-z ]+', '', text)  
  
    # menghapus spasi berlebih dan mengurangi [semua] kata  
    # spasi ganda diubah menjadi satu spasi saja  
    text = re.sub(' +', ' ', text.strip())  
  
    # mengembalikan text yang telah melalui proses case folding  
    return text  
  
# menerapkan fungsi cleansing pada kolom 'komentar'  
# sehingga disimpan pada kolom baru bernama 'clean_text'  
df['clean_text'] = df['komentar'].apply(cleansing)  
  
# mengatur agar output text pada kolom 'clean_text' secara pandas  
pd.set_option('display.max_columns', None)  
  
# menampilkan 10 baris pertama  
print(df)
```

Gambar 1. *Cleansing Data*

Berikut contoh hasil ulasan sebelum dan sesudah melalui proses pembersihan yang terdapat pada Tabel 1.

Tabel 1. Tabel hasil cleansing

| no | komentar                                                                                                               | clean_text                                                                                                          |
|----|------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|
| 1  | yakin mereka gak bawa motor? sy kok curiga ada sewa parkir dadakan                                                     | yakin mereka gak bawa motor sy kok curiga ada sewa parkir dadakan                                                   |
| 2  | Keren, wajib berlaku di seluruh Indonesia ini                                                                          | keren wajib berlaku di seluruh indonesia ini                                                                        |
| 3  | Yg 11 km dari sekolah                                                                                                  | yg km dari sekolah                                                                                                  |
| 4  | Gubernur jkt lagi sibuk mindahin patung dan buat patung baru,sambil ngejar ngejar penunggak pajak kendaraan bermotor?? | gubernur jkt lagi sibuk mindahin patung dan buat patung barusambil ngejar ngejar penunggak pajak kendaraan bermotor |

|    |                                                                                                                                                                                                    |                                                                                                                                                                                  |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 5  | Sangat setuju banget sama sangat cocok bgt ya.... Lbh baik satu tahun aja ya..... Bravo kang Dedi Mulyadi                                                                                          | sangat setuju banget sama sangat cocok bgt ya lbh baik satu tahun aja ya bravo kang dedi mulyadi                                                                                 |
| 6  | dari umur berapa???                                                                                                                                                                                | dari umur berapa                                                                                                                                                                 |
| 7  | mau jam brp GK jadi masalah yg jadi masalah pendidikan GK gratis di Malaysia swasta ja gratis 10 tahun Jokowi bikin pendidikan ketinggalan jauh dr negara Malaysia                                 | mau jam brp gk jadi masalah yg jadi masalah pendidikan gk gratis di malaysia swasta ja gratis tahun Jokowi bikin pendidikan ketinggalan jauh dr negara malaysia                  |
| 8  | masuk jam 6.30 dah di anggap beban gua di jakarta aja masuk jam 6.00 pulang jam 16.00 ga jadi beban, karna udah biasa dan ga males malesan itu juga buat belajar di siplin dengan waktu, bagus itu | masuk jam dah di anggap beban gua di jakarta aja masuk jam pulang jam ga jadi beban karna udah biasa dan ga males malesan itu juga buat belajar di siplin dengan waktu bagus itu |
| 9  | Kenapa orang dewasa suka terobsesi menerapkan standar dirinya pada anak <sup>2</sup> . Kenapa tidak memperhitungkan hak anak. Mentang <sup>2</sup> anak <sup>2</sup> gak punya kekuatan            | kenapa orang dewasa suka terobsesi menerapkan standar dirinya pada anak kenapa tidak memperhitungkan hak anak mentang anak gak punya kekuatan                                    |
| 10 | Ga ngaruh di sekolahku tetep jam 13.15                                                                                                                                                             | ga ngaruh di sekolahku tetep jam                                                                                                                                                 |
| 11 | 6.30 normal kali                                                                                                                                                                                   | normal kali                                                                                                                                                                      |
| 12 | setujuuuuuu banget!!!!                                                                                                                                                                             | setujuuuuuu banget                                                                                                                                                               |
| 13 | cirebon banyak bagt,,                                                                                                                                                                              | cirebon banyak bagt                                                                                                                                                              |
| 14 | banjir we nu di bahas teh itu hukum alam banjir mah                                                                                                                                                | banjir we nu di bahas teh itu hukum alam banjir mah                                                                                                                              |
| 15 | banyak yg protes pak karena istilah tidak berbau agama coba kalau pake kata sedekah                                                                                                                | banyak yg protes pak karena istilah tidak berbau agama coba kalau pake kata sedekah                                                                                              |
| 16 | mantap gubernur Jabar                                                                                                                                                                              | mantap gubernur jabar                                                                                                                                                            |
| 17 | udah pajak seglanya udah di potong langsung ngapain juga                                                                                                                                           | udah pajak seglanya udah di potong langsung ngapain juga                                                                                                                         |

## 2) Normalisasi

Tahap normalisasi teks bertujuan menyeragamkan kata tidak baku, *slangword*, dan singkatan menjadi kata baku agar kualitas data meningkat serta mengurangi variasi penulisan yang memiliki makna sama. Proses diawali dengan mengimpor pustaka **pandas** dan membaca kamus normalisasi dari berkas **lexicon.csv**, kemudian mengubah seluruh kata menjadi huruf kecil dan menyimpannya dalam bentuk *dictionary*. Kamus tersebut digabungkan dengan kamus tambahan untuk memperkaya kosakata. Selanjutnya, setiap kata pada kolom **clean\_text** diperiksa dan diubah menjadi bentuk baku sesuai kamus, sedangkan kata yang tidak ditemukan tetap dipertahankan. Hasil normalisasi disimpan pada kolom **normalized** dan diverifikasi sebelum digunakan pada tahap berikutnya.

Berikut contoh hasil ulasan sebelum dan sesudah melalui proses normalization yang terdapat pada Tabel 2.

**Tabel 2.** Tabel hasil Normalisasi

| no | komentar                                                                          | normalized                                                                           |
|----|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| 1  | yakin mereka ngak bawa motor? sy kok curiga ada sewa parkir dadakan               | yakin mereka enggak bawa motor saya kok curiga ada sewa parkir dadakan               |
| 2  | Keren, wajib berlaku di seluruh Indonesia ini                                     | keren wajib berlaku di seluruh indonesia ini                                         |
| 3  | Yg 11 km dari sekolah                                                             | yang kamu dari sekolah                                                               |
| 4  | Gubernur jkt lagi sibuk mindahin patung dan buat patung baru,sambil ngejar ngejar | gubernur jakarta lagi sibuk mindahin patung dan buat patung barusambil ngejar ngejar |

|    |                                                                                                                                                                                                    |                                                                                                                                                                                           |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    | penunggak pajak kendaraan bermotor??                                                                                                                                                               | penunggak pajak kendaraan bermotor                                                                                                                                                        |
| 5  | Sangat setuju banget sama sangat cocok bgt ya.... Lbh baik satu tahun aja ya..... Bravo kang Dedi Mulyadi                                                                                          | sangat setuju banget sama sangat cocok banget ya lebih baik satu tahun saja ya bravo kang dedi mulyadi                                                                                    |
| 6  | dari umur berapa???                                                                                                                                                                                | dari umur berapa                                                                                                                                                                          |
| 7  | mau jam brp GK jadi masalah yg jadi masalah pendidikan GK gratis di Malaysia swasta ja gratis 10 tahun Jokowi bikin pendidikan ketinggalan jauh dr negara Malaysia                                 | mau jam berapa tidak jadi masalah yang jadi masalah pendidikan tidak gratis di malaysia swasta saja gratis tahun jokowi bikin pendidikan ketinggalan jauh dari negara malaysia            |
| 8  | masuk jam 6.30 dah di anggap beban gua di jakarta aja masuk jam 6.00 pulang jam 16.00 ga jadi beban, karna udah biasa dan ga males malesan itu juga buat belajar di siplin dengan waktu, bagus itu | masuk jam deh di anggap beban gua di jakarta saja masuk jam pulang jam tidak jadi beban karena sudah biasa dan tidak malas malesan itu juga buat belajar di siplin dengan waktu bagus itu |
| 9  | Kenapa orang dewasa suka terobsesi menerapkan standar dirinya pada anak <sup>2</sup> . Kenapa tidak memperhitungkan hak anak. Mentang <sup>2</sup> anak <sup>2</sup> gak punya kekuatan            | kenapa orang dewasa suka terobsesi menerapkan standar dirinya pada anak kenapa tidak memperhitungkan hak anak mentang anak enggak punya kekuatan                                          |
| 10 | Ga ngaruh di sekolahku tetep jam 13.15                                                                                                                                                             | tidak ngaruh di sekolahku tetap jam                                                                                                                                                       |
| 11 | 6.30 normal kali                                                                                                                                                                                   | normal kali                                                                                                                                                                               |
| 12 | setujuuuuuu banget!!!!                                                                                                                                                                             | setujuuuuuu banget                                                                                                                                                                        |
| 13 | cirebon banyak bagt,,                                                                                                                                                                              | cirebon banyak bagt                                                                                                                                                                       |
| 14 | banjir we nu di bahas teh itu hukum alam banjir mah                                                                                                                                                | banjir we nu di bahas teh itu hukum alam banjir mah                                                                                                                                       |
| 15 | banyak yg protes pak karena istilah tidak berbau agama coba kalau pake kata sedekah                                                                                                                | banyak yang protes pak karena istilah tidak berbau agama coba kalau pakai kata sedekah                                                                                                    |
| 16 | mantap gubernur Jabar                                                                                                                                                                              | mantap gubernur jabar                                                                                                                                                                     |
| 17 | udah pajak seglanya udah di potong langsung ngapain juga                                                                                                                                           | sudah pajak seglanya sudah di potong langsung mengapai juga                                                                                                                               |

### 3) Stopword Removal

Tahap *stopword removal* bertujuan menghapus kata-kata umum yang tidak berkontribusi terhadap penentuan sentimen sehingga dapat mengurangi *noise* dan meningkatkan kualitas data. Proses ini menggunakan kombinasi kamus *stopword* bawaan pustaka **Sastrawi** dan kamus *stopword* kustom berisi sekitar 200 kata yang disesuaikan dengan karakteristik bahasa media sosial. Kata-kata negasi seperti *tidak*, *bukan*, *belum*, *kurang*, *tanpa*, dan *jangan* tetap dipertahankan karena berpengaruh terhadap polaritas sentimen. Proses dilakukan dengan memuat kedua kamus, menggabungkannya, menerapkan fungsi *stopword removal* pada kolom **normalized**, kemudian menyimpan hasilnya pada kolom **clean\_no\_stop** sebagai data yang lebih ringkas dan siap digunakan dalam pemodelan **Long Short-Term Memory (LSTM)**.

Berikut contoh hasil ulasan sebelum dan sesudah melalui proses *stopword removal* yang terdapat pada Tabel 3.

**Tabel 3.** Tabel hasil Stopword removal

| no | normalized                                                                                                              | clean_no_stop                                                                                                  |
|----|-------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| 1  | yakin mereka enggak bawa motor saya kok curiga ada sewa parkir dadakan                                                  | yakin enggak bawa motor curiga sewa parkir dadakan                                                             |
| 2  | keren wajib berlaku di seluruh indonesia ini                                                                            | keren wajib berlaku seluruh indonesia                                                                          |
| 3  | yang kamu dari sekolah                                                                                                  | kamu sekolah                                                                                                   |
| 4  | gubernur jakarta lagi sibuk mindahin patung dan buat patung barusambil ngejar ngejar penunggak pajak kendaraan bermotor | gubernur jakarta sibuk mindahin patung buat patung barusambil ngejar ngejar penunggak pajak kendaraan bermotor |
| 5  | sangat setuju banget sama sangat cocok                                                                                  | setuju sama cocok lebih baik satu bravo kang dedi mulyadi                                                      |

|    |                                                                                                                                                                                           |                                                                                                                                        |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
|    | banget ya lebih baik satu tahun saja ya bravo kang dedi mulyadi                                                                                                                           |                                                                                                                                        |
| 6  | dari umur berapa                                                                                                                                                                          | umur berapa                                                                                                                            |
| 7  | mau jam berapa tidak jadi masalah yang jadi masalah pendidikan tidak gratis di malaysia swasta gratis tahun jokowi bikin pendidikan ketinggalan jauh dari negara malaysia                 | mau jam berapa tidak jadi jadi pendidikan tidak gratis malaysia swasta gratis jokowi bikin pendidikan ketinggalan jauh negara malaysia |
| 8  | masuk jam deh di anggap beban gua di jakarta saja masuk jam pulang jam tidak jadi beban karena sudah biasa dan tidak malas malesan itu juga buat belajar di siplin dengan waktu bagus itu | masuk jam anggap beban gua jakarta masuk jam pulang jam tidak jadi beban biasa tidak malas malesan buat belajar siplin waktu bagus     |
| 9  | kenapa orang dewasa suka terobsesi menerapkan standar dirinya pada anak kenapa tidak memperhitungkan hak anak mentang anak enggak punya kekuatan                                          | orang dewasa suka terobsesi menerapkan standar dirinya anak tidak memperhitungkan hak anak mentang anak enggak punya kekuatan          |
| 10 | tidak ngaruh di sekolahku tetap jam                                                                                                                                                       | tidak ngaruh sekolahku tetap jam                                                                                                       |
| 11 | normal kali                                                                                                                                                                               | normal kali                                                                                                                            |
| 12 | setujuuuuuu banget                                                                                                                                                                        | setujuuuuuu                                                                                                                            |
| 13 | cirebon banyak bagt                                                                                                                                                                       | cirebon banyak bagt                                                                                                                    |
| 14 | banjir we nu di bahas teh itu hukum alam banjir mah                                                                                                                                       | banjir we nu bahas teh hukum alam banjir mah                                                                                           |
| 15 | banyak yang protes pak karena istilah tidak berbau agama coba kalau pakai kata sedekah                                                                                                    | banyak protes istilah tidak berbau agama coba kalau pakai kata sedekah                                                                 |
| 16 | mantap gubernur jabar                                                                                                                                                                     | mantap gubernur jabar                                                                                                                  |
| 17 | sudah pajak seglanya sudah di potong langsung mengapai juga                                                                                                                               | pajak seglanya potong langsung mengapai                                                                                                |

#### 4) Stemming

Tahap *stemming* bertujuan mengubah kata berimbuhan menjadi kata dasar untuk mengurangi variasi kata dan meningkatkan konsistensi data. Proses ini menggunakan pustaka **Sastrawi** dengan menerapkan fungsi *stemming* pada kolom **clean\_no\_stop**, kemudian menyimpan hasilnya pada **clean\_text** sebagai data akhir prapemrosesan untuk mendukung kinerja model **Long Short-Term Memory (LSTM)**.

Berikut contoh hasil ulasan sebelum dan sesudah melalui proses stemming yang terdapat pada Tabel 4

**Tabel 4.** Tabel hasil Stemming

| no | Stop removal                                                                                                                           | stemming                                                                                                                 |
|----|----------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| 1  | yakin enggak bawa motor curiga sewa parkir dadakan                                                                                     | yakin enggak bawa motor curiga sewa parkir dada                                                                          |
| 2  | keren wajib berlaku seluruh indonesia                                                                                                  | keren wajib laku seluruh indonesia                                                                                       |
| 3  | kamu sekolah                                                                                                                           | kamu sekolah                                                                                                             |
| 4  | gubernur jakarta sibuk mindahin patung buat patung barusambil ngejar ngejar penunggak pajak kendaraan bermotor                         | gubernur jakarta sibuk mindahin patung buat patung barusambil ngejar ngejar tunggak pajak kendra motor                   |
| 5  | setuju sama cocok lebih baik satu bravo kang dedi mulyadi                                                                              | tuju sama cocok lebih baik satu bravo kang dedi mulyadi                                                                  |
| 6  | umur berapa                                                                                                                            | umur berapa                                                                                                              |
| 7  | mau jam berapa tidak jadi jadi pendidikan tidak gratis malaysia swasta gratis jokowi bikin pendidikan ketinggalan jauh negara malaysia | mau jam berapa tidak jadi jadi didik tidak gratis malaysia swasta gratis jokowi bikin didik tinggal jauh negara malaysia |
| 8  | masuk jam anggap beban gua jakarta masuk jam pulang jam tidak jadi beban biasa tidak malas                                             | masuk jam anggap beban gua jakarta masuk jam pulang jam tidak jadi beban                                                 |

|    |                                                                                                                               |                                                                                                       |
|----|-------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------|
|    | malesan buat belajar siplin waktu bagus                                                                                       | biasa tidak malas malesan buat ajar siplin waktu bagus                                                |
| 9  | orang dewasa suka terobsesi menerapkan standar dirinya anak tidak memperhitungkan hak anak mentang anak enggak punya kekuatan | orang dewasa suka obsesi terap standar diri anak tidak hitung hak anak mentang anak enggak punya kuat |
| 10 | tidak ngaruh sekolahku tetap jam                                                                                              | tidak ngaruh sekolah tetap jam                                                                        |
| 11 | normal kali                                                                                                                   | normal kali                                                                                           |
| 12 | setujuuuuuu                                                                                                                   | setujuuuuuu                                                                                           |
| 13 | cirebon banyak bagt                                                                                                           | cirebon banyak bagt                                                                                   |
| 14 | banjir we nu bahas teh hukum alam banjir mah                                                                                  | banjir we nu bahas teh hukum alam banjir mah                                                          |
| 15 | banyak protes istilah tidak berbau agama coba kalau pakai kata sedekah                                                        | banyak protes istilah tidak bau agama coba kalau pakai kata sedekah                                   |
| 16 | mantap gubernur jabar                                                                                                         | mantap gubernur jabar                                                                                 |
| 17 | pajak seglanya potong langsung mengapai                                                                                       | pajak seglanya potong langsung kapai                                                                  |

#### 5) Tokenization

Tahap tokenisasi bertujuan memecah teks hasil prapemrosesan menjadi token berupa kata menggunakan metode *whitespace tokenization*. Proses ini diterapkan pada kolom **clean\_text**, kemudian hasilnya disimpan pada kolom **tokens**. Token yang dihasilkan menjadi representasi numerik awal yang digunakan dalam pemodelan sentimen menggunakan algoritma **Long Short-Term Memory (LSTM)**.

Berikut hasil ulasan sebelum dan sesudah melalui proses tokenisasi yang terdapat pada Tabel 5.

**Tabel 5.** Tabel hasil tokenisasi

| no | tokens                                                                                                                                                                                             |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | ['yakin', 'enggak', 'bawa', 'motor', 'curiga', 'sewa', 'parkir', 'dada']                                                                                                                           |
| 2  | ['keren', 'wajib', 'laku', 'seluruh', 'indonesia']                                                                                                                                                 |
| 3  | ['kamu', 'sekolah']                                                                                                                                                                                |
| 4  | ['gubernur', 'jakarta', 'sibuk', 'mindahin', 'patung', 'buat', 'patung', 'barusambil', 'ngejar', 'ngejar', 'tunggak', 'pajak', 'kendara', 'motor']                                                 |
| 5  | ['tjuu', 'sama', 'cocok', 'lebih', 'baik', 'satu', 'bravo', 'kang', 'dedi', 'mulyadi']                                                                                                             |
| 6  | ['umur', 'berapa']                                                                                                                                                                                 |
| 7  | ['mau', 'jam', 'berapa', 'tidak', 'jadi', 'jadi', 'didik', 'tidak', 'gratis', 'malaysia', 'swasta', 'gratis', 'jokowi', 'bikin', 'didik', 'tinggal', 'jauh', 'negara', 'malaysia']                 |
| 8  | ['masuk', 'jam', 'anggap', 'beban', 'gua', 'jakarta', 'masuk', 'jam', 'pulang', 'jam', 'tidak', 'jadi', 'beban', 'biasa', 'tidak', 'malas', 'malesan', 'buat', 'ajar', 'siplin', 'waktu', 'bagus'] |
| 9  | ['orang', 'dewasa', 'suka', 'obsesi', 'terap', 'standar', 'diri', 'anak', 'tidak', 'hitung', 'hak', 'anak', 'mentang', 'anak', 'enggak', 'punya', 'kuat']                                          |
| 10 | ['tidak', 'ngaruh', 'sekolah', 'tetap', 'jam']                                                                                                                                                     |
| 11 | ['normal', 'kali']                                                                                                                                                                                 |
| 12 | ['setujuuuuuu']                                                                                                                                                                                    |
| 13 | ['cirebon', 'banyak', 'bagt']                                                                                                                                                                      |
| 14 | ['banjir', 'we', 'nu', 'bahas', 'teh', 'hukum', 'alam', 'banjir', 'mah']                                                                                                                           |
| 15 | ['banyak', 'protes', 'istilah', 'tidak', 'bau', 'agama', 'coba', 'kalau', 'pakai', 'kata', 'sedekah']                                                                                              |
| 16 | ['mantap', 'gubernur', 'jabar']                                                                                                                                                                    |
| 17 | ['pajak', 'seglanya', 'potong', 'langsung', 'kapai']                                                                                                                                               |

#### d. Pelabelan Data

Pada tahap pelabelan, setiap ulasan diklasifikasikan ke dalam sentimen positif, negatif, atau netral menggunakan pendekatan Lexicon-Based berbasis kamus sentimen KBBI. Proses dilakukan dengan mencocokkan setiap kata pada hasil *preprocessing* dengan kamus *lexicon*, menghitung skor sentimen, kemudian menentukan label

berdasarkan nilai skor. Hasil pelabelan disimpan pada kolom *lexicon\_score* dan *label\_lexicon*, sementara label manual tetap dipertahankan sebagai *ground truth* untuk menjaga keakuratan data.

Dengan pendekatan ini, pelabelan sentimen dapat dilakukan secara sistematis tanpa mengubah label manual yang telah dilabeli oleh manusia, sehingga seluruh data dapat dimanfaatkan secara optimal dalam proses analisis sentiment.

**Tabel 6.** Tabel Hasil Labeling

| <b>clean_no_stop</b>                                                                                                                                        | <b>lexicon_score</b> | <b>label_lexicon</b> |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|----------------------|
| kalau bandung enak udara adem                                                                                                                               | 0                    | netral               |
| jawa barat gebrak nya jawa tengah gebrak nya mana                                                                                                           | 0                    | netral               |
| tuju pis coooo                                                                                                                                              | 1                    | positif              |
| nah gin aku tuju apaya kata seragam guna menyamakantidak beda kasta ekonomi nya kalo pake motor liat kan untung pak dedi masa iya satu seragam beda kendara | 1                    | positif              |
| aku biyen baru beli motor ama ortu pas kuliah semester sd sma mancangess salam era                                                                          | 0                    | netral               |
| padahall pakai sepeda lebih ramah lingkung yah                                                                                                              | 0                    | netral               |
| kalau gak naik sepeda aja biar kayak dijepang                                                                                                               | 0                    | netral               |
| seru banget ingat sekolah dulu                                                                                                                              | 0                    | netral               |
| negara maju aja anak sekolah gitu                                                                                                                           | 1                    | positif              |
| kereennnnn kyak zaman ku dulu tahun an mantappp bakam kamisam                                                                                               | 0                    | netral               |
| karna baca nama dedy muliyadi ge kbayang nyemplung kolam ikan                                                                                               | 0                    | netral               |
| bagusbpkaink                                                                                                                                                | 0                    | netral               |
| banyak polisi yg jaga jalur yg rawan tawur ajar                                                                                                             | 0                    | netral               |
| luar biasa                                                                                                                                                  | 0                    | netral               |
| lebih asikkkkk                                                                                                                                              | 0                    | netral               |
| seruu                                                                                                                                                       | 0                    | netral               |
| dulu aku teman pencet bel rumah orangorang kaya panjang jalan tuju pangkal angkot                                                                           | 1                    | positif              |
| moga ikut daerah lainlanjutan pak                                                                                                                           | 0                    | netral               |
| ntr banyak yg nebeng mobil truk                                                                                                                             | 0                    | netral               |

Data hasil pelabelan yang telah diproses kemudian disimpan ke dalam sebuah berkas berformat CSV dengan nama pelabelan.csv. File ini berisi teks hasil *preprocessing* (*clean\_no\_stop*) beserta label sentimen hasil pelabelan lexicon. Penyimpanan data dilakukan agar hasil pelabelan dapat digunakan kembali secara konsisten pada tahap selanjutnya.

Berdasarkan hasil pelabelan sentimen menggunakan metode *lexicon-based* dengan pembobotan kata, diperoleh distribusi sentimen terhadap data yang dianalisis. Dari total 10.751 dataset, setelah dilakukan proses pelabelan, diperoleh 6.440 data yang diklasifikasikan sebagai sentimen netral, 1.944 data sebagai sentimen positif dan 1.944 data sebagai sentimen negatif.

#### **e. Pembobotan kata (Word Embedding)**

Pada penelitian ini, *word embedding* digunakan untuk mengubah kata menjadi vektor numerik berdimensi tetap agar dapat diproses oleh model Long Short-Term Memory (LSTM). Proses dimulai dengan tokenisasi menggunakan Tokenizer TensorFlow, dilanjutkan dengan *padding* menggunakan *pad\_sequences* agar panjang setiap teks seragam. Selanjutnya dibentuk embedding layer berdimensi 128 yang memetakan setiap indeks kata menjadi representasi vektor. Vektor embedding kemudian diekstraksi untuk menganalisis hubungan semantik antar kata. Representasi ini memungkinkan kata yang memiliki konteks serupa berada pada posisi vektor yang berdekatan sehingga meningkatkan kemampuan model LSTM dalam mengklasifikasikan sentimen. Contoh hasil *word embedding* disajikan pada Tabel 7.

**Tabel 7.** Tabel Hasil Pembobotan kata

DOI: <https://doi.org/10.31004/riggs.v5i2.11497>  
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

| kata    | index | dim_1    | dim_2    | dim_3    | dim_4    | dim_5    |
|---------|-------|----------|----------|----------|----------|----------|
| anak    | 3     | -0,02496 | 0,049621 | -0,04413 | -0,01154 | 0,024118 |
| jam     | 4     | -0,00422 | -0,00236 | -0,02781 | -0,00158 | 0,023551 |
| sekolah | 5     | -0,0393  | -0,00296 | -0,01349 | 0,047896 | 0,038199 |
| pak     | 6     | 0,033629 | -0,02104 | 0,00087  | 0,044044 | -0,02218 |
| jadi    | 8     | -0,03606 | -0,04238 | -0,02656 | -0,01865 | 0,036616 |
| tuju    | 9     | -0,02051 | -0,04535 | 0,036935 | 0,021714 | 0,02851  |
| masuk   | 10    | -0,04383 | 0,016996 | -0,01676 | -0,03388 | 0,029195 |
| dulu    | 13    | -0,04787 | -0,02783 | -0,02052 | 0,002038 | 0,023283 |
| banyak  | 14    | 0,025404 | 0,016802 | 0,035705 | -0,03135 | -0,02469 |
| sama    | 15    | 0,047838 | -0,01241 | -0,02701 | -0,01469 | 0,011347 |
| jalan   | 16    | 0,02228  | -0,03039 | 0,045273 | 0,045237 | -0,0047  |
| orang   | 17    | -0,03776 | -0,02957 | -0,00492 | 0,008803 | 0,016486 |
| lebih   | 18    | 0,007176 | -0,00087 | 0,031498 | 0,015082 | -0,00864 |
| buat    | 19    | -0,01196 | -0,0169  | 0,043568 | 0,041209 | -0,00745 |
| sudah   | 20    | -0,01869 | -0,03818 | -0,04139 | 0,003777 | 0,027661 |

Kata-kata yang ditampilkan pada tabel merupakan bagian dari vocabulary tokenizer, yaitu kumpulan kata unik yang terbentuk setelah melalui tahapan *preprocessing* teks, meliputi *case folding*, normalisasi, penghapusan *stopword*, dan *stemming*. Oleh karena itu, kata yang muncul pada tabel adalah kata dalam bentuk dasar hasil *preprocessing*, bukan kata mentah dari data asli.

Nilai pada setiap dimensi embedding tidak memiliki makna secara individual, namun secara kolektif membentuk representasi semantik suatu kata. Kata-kata yang memiliki konteks dan makna yang serupa dalam kalimat cenderung memiliki pola vektor embedding yang mendekati satu sama lain. Representasi ini memungkinkan model LSTM untuk menangkap hubungan semantik antar kata secara lebih efektif dalam proses klasifikasi sentimen.

#### f. Pemodelan (LSTM)

Tahap pemodelan bertujuan memperoleh performa dan akurasi model Long Short-Term Memory (LSTM) dalam klasifikasi sentimen. Proses diawali dengan mengimpor pustaka pendukung, kemudian mengubah data teks menjadi representasi numerik menggunakan Tokenizer dan padding. Dataset dibagi menjadi data latih dan data uji, sedangkan label sentimen dikonversi ke bentuk numerik menggunakan LabelEncoder. Model dibangun dengan arsitektur Sequential yang terdiri atas Embedding, LSTM, Dense, Dropout, dan Softmax. Pelatihan menggunakan optimizer Adam, fungsi *loss* sparse categorical crossentropy, serta *class weight* untuk mengatasi ketidakseimbangan data. Hasil evaluasi akurasi pada berbagai rasio pembagian data disajikan pada Tabel 8.

**Tabel 8.** Tabel Rasio Perbandingan Akurasi

| no | Rasio Perbandingan | data latih | data uji | Akurasi |
|----|--------------------|------------|----------|---------|
| 1  | 90:10              | 9728       | 1081     | 81,04%  |
| 2  | 80:20              | 8647       | 2162     | 82,79%  |
| 3  | 70:30              | 7566       | 3243     | 75,89%  |
| 4  | 60:40              | 6485       | 4324     | 79,65%  |

Proses pelatihan dan pengujian model klasifikasi sentimen menggunakan metode Long Short-Term Memory (LSTM) dilakukan dengan menerapkan beberapa skenario pembagian data, yaitu rasio 90:10, 80:20, 70:30, dan 60:40 antara data latih dan data uji. Evaluasi performa model dilakukan berdasarkan nilai akurasi pada data uji (*validation accuracy*) untuk menilai kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya.

Pada skenario pembagian data 90:10, model LSTM memperoleh akurasi data uji sebesar 81,04%. Hasil ini menunjukkan bahwa dengan proporsi data latih yang lebih besar, model mampu mempelajari pola sentimen secara efektif dan menghasilkan performa yang cukup baik pada data uji.

Pada skenario pembagian data 80:20, model menunjukkan performa terbaik dengan nilai akurasi data uji tertinggi, yaitu sebesar 82,79%. Hasil ini mengindikasikan bahwa rasio 80:20 memberikan keseimbangan yang optimal antara jumlah data latih dan data uji, sehingga model mampu melakukan generalisasi dengan lebih baik.

Selanjutnya, pada skenario pembagian data 70:30, akurasi data uji yang diperoleh menurun menjadi 79,65%. Penurunan ini disebabkan oleh berkurangnya jumlah data latih yang digunakan untuk membangun model, sehingga kemampuan model dalam menangkap pola sentimen menjadi sedikit berkurang.

Pada skenario pembagian data 60:40, model memperoleh akurasi data uji sebesar 75,89%, yang merupakan nilai terendah dibandingkan skenario lainnya. Hal ini menunjukkan bahwa proporsi data uji yang terlalu besar dapat mempengaruhi kinerja model karena keterbatasan data latih dalam proses pembelajaran.

Secara keseluruhan, hasil pengujian menunjukkan bahwa model LSTM memiliki performa yang cukup stabil dengan akurasi data uji berada di atas 75% pada seluruh skenario pembagian data. Rasio 80:20 menghasilkan performa terbaik dengan akurasi tertinggi sebesar 82,79%, sehingga rasio tersebut dipilih sebagai konfigurasi optimal untuk model klasifikasi sentimen menggunakan metode LSTM dalam penelitian ini.

### g. Evaluasi Kinerja Model

Tahap evaluasi dilakukan untuk mengetahui seberapa baik kinerja model *Long Short-Term Memory* (LSTM) dalam mengklasifikasikan sentimen pada data uji. Evaluasi ini bertujuan untuk mengukur kemampuan model dalam melakukan prediksi terhadap data yang belum pernah digunakan pada tahap pelatihan.

#### 1) Hasil Evaluasi Kinerja Model

Pada tahap awal, dilakukan proses prediksi terhadap data uji menggunakan model LSTM yang telah dilatih. Model menghasilkan keluaran berupa probabilitas untuk setiap kelas sentimen. Nilai probabilitas tersebut kemudian dikonversi menjadi label kelas dengan memilih indeks kelas dengan nilai probabilitas tertinggi menggunakan fungsi `argmax`.

Berikut hasil Perhitungan Model prediksi pada table 9

**Tabel 9.** Tabel Hasil Nilai Probabilitas Model

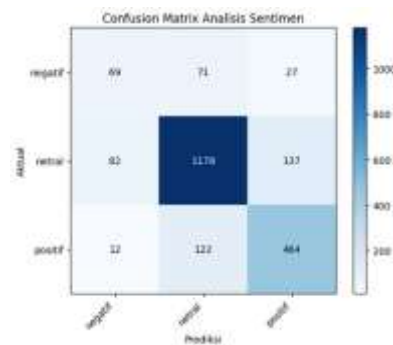
| komentar                                                                                                                                                                                           | prob_negatif | prob_netral | prob_positif |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-------------|--------------|
| yakin mereka gak bawa motor? sy kok curiga ada sewa parkir dadakan                                                                                                                                 | 0,352458954  | 0,321608126 | 0,32593289   |
| Keren, wajib berlaku di seluruh Indonesia ini                                                                                                                                                      | 0,336705506  | 0,308552206 | 0,354742318  |
| Yg 11 km dari sekolah                                                                                                                                                                              | 0,350009531  | 0,311737895 | 0,338252544  |
| Gubernur jkt lagi sibuk mindahin patung dan buat patung baru,sambil ngejar ngejar penunggak pajak kendaraan bermotor??                                                                             | 0,339023858  | 0,330405354 | 0,330570847  |
| Sangat setuju banget sama sangat cocok bgt ya.... Lbh baik satu tahun aja ya..... Bravo kang Dedi Mulyadi                                                                                          | 0,340091437  | 0,324657083 | 0,335251451  |
| dari umur berapa???                                                                                                                                                                                | 0,341421425  | 0,315441042 | 0,343137413  |
| mau jam brp GK jadi masalah yg jadi masalah pendidikan GK gratis di Malaysia swasta ja gratis 10 tahun Jokowi bikin pendidikan ketinggalan jauh dr negara Malaysia                                 | 0,352795094  | 0,327153504 | 0,320051372  |
| masuk jam 6.30 dah di anggap beban gua di jakarta aja masuk jam 6.00 pulang jam 16.00 ga jadi beban, karna udah biasa dan ga males malesan itu juga buat belajar di siplin dengan waktu, bagus itu | 0,330912799  | 0,326916963 | 0,342170268  |
| Kenapa orang dewasa suka terobsesi                                                                                                                                                                 | 0,339385003  | 0,326890022 | 0,333725065  |

|                                                                                                                                                            |             |             |             |
|------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-------------|-------------|
| menerapkan standar dirinya pada anak <sup>2</sup> .<br>Kenapa tidak memperhitungkan hak anak.<br>Mentang <sup>2</sup> anak <sup>2</sup> gak punya kekuatan |             |             |             |
| Ga ngaruh di sekolahku tetep jam 13.15                                                                                                                     | 0,355708957 | 0,311079115 | 0,333211958 |
| 6.30 normal kali                                                                                                                                           | 0,339471787 | 0,306459606 | 0,354068637 |
| setujuuuuuu banget!!!!                                                                                                                                     | 0,344180882 | 0,305618197 | 0,350200891 |
| cirebon banyak bagt,,                                                                                                                                      | 0,353785962 | 0,30166164  | 0,344552398 |
| banjir we nu di bahas teh itu hukum alam<br>banjir mah                                                                                                     | 0,331864834 | 0,351629734 | 0,316505373 |
| banyak yg protes pak karena istilah tidak<br>berbau agama coba kalau pake kata sedekah                                                                     | 0,325316727 | 0,356802523 | 0,317880601 |
| mantap gubernur Jabar                                                                                                                                      | 0,33899048  | 0,306199074 | 0,354810327 |
| udah pajak seglanya udah di potong langsung<br>ngapain juga                                                                                                | 0,340606749 | 0,316918969 | 0,342474312 |

Selanjutnya, dilakukan perhitungan nilai akurasi menggunakan fungsi *accuracy\_score* dari pustaka *scikit-learn*. Nilai akurasi digunakan untuk mengukur persentase prediksi yang benar dibandingkan dengan keseluruhan data uji.

Selain akurasi, evaluasi kinerja model juga dilakukan menggunakan *classification report*, yang mencakup nilai *precision*, *recall*, dan *f1-score* untuk setiap kelas sentimen, yaitu negatif, netral, dan positif. Nilai *precision* menunjukkan tingkat ketepatan model dalam memprediksi suatu kelas, *recall* menunjukkan kemampuan model dalam mengenali seluruh data aktual pada kelas tersebut, sedangkan *f1-score* merupakan rata-rata harmonis antara *precision* dan *recall*.

Hasil *classification report* memberikan gambaran yang lebih rinci mengenai performa model pada masing-masing kelas sentimen, sehingga dapat diketahui kelas mana yang paling baik maupun paling sulit diklasifikasikan oleh model LSTM. Evaluasi ini menjadi dasar dalam menilai kualitas model secara menyeluruh, tidak hanya berdasarkan akurasi, tetapi juga dari keseimbangan performa antar kelas.



**Gambar 2.** Hasil Evaluasi Confusion Matrix

Hasil evaluasi kinerja model klasifikasi sentimen menggunakan metode *Long Short-Term Memory* (LSTM) ditunjukkan melalui nilai *accuracy*, *precision*, *recall*, dan *f1-score* yang diperoleh dari data uji sebanyak 2.162 data. Berdasarkan hasil pengujian, model menghasilkan akurasi sebesar 79%, yang menunjukkan bahwa sebagian besar data uji berhasil diklasifikasikan dengan benar oleh model.

Pada kelas sentimen negatif, model memperoleh nilai *precision* sebesar 0,42, *recall* sebesar 0,41, dan *f1-score* sebesar 0,42 dengan jumlah data sebanyak 167. Nilai ini menunjukkan bahwa model masih mengalami kesulitan dalam mengenali sentimen negatif secara konsisten, yang disebabkan oleh jumlah data negatif yang relatif lebih sedikit dibandingkan kelas lainnya serta adanya kemiripan pola kata dengan kelas netral.

Untuk kelas sentimen netral, model menunjukkan performa terbaik dengan nilai *precision* sebesar 0,86, *recall* sebesar 0,84, dan *f1-score* sebesar 0,85 dari total 1.397 data. Hasil ini mengindikasikan bahwa model LSTM sangat baik dalam mengidentifikasi sentimen netral, yang juga merupakan kelas dominan dalam dataset.

Sementara itu, pada kelas sentimen positif, model memperoleh nilai precision sebesar 0,74, recall sebesar 0,78, dan f1-score sebesar 0,76 dengan jumlah data sebanyak 598. Nilai recall yang lebih tinggi dibandingkan precision menunjukkan bahwa model cukup baik dalam mengenali data positif, meskipun masih terdapat beberapa data positif yang salah diklasifikasikan sebagai sentimen netral.

Berdasarkan nilai weighted average, model menghasilkan precision, recall, dan f1-score masing-masing sebesar 0,79, yang menunjukkan bahwa secara keseluruhan model LSTM memiliki performa yang cukup baik dan stabil dalam mengklasifikasikan sentimen pada data uji.

Hasil evaluasi ini menunjukkan bahwa model LSTM mampu melakukan klasifikasi sentimen dengan baik, terutama pada kelas netral dan positif. Namun, performa pada kelas negatif masih perlu ditingkatkan, misalnya dengan penyeimbangan data (class balancing) atau penambahan data latih pada kelas negatif.

**Tabel 10.** hasil klasifikasi model

| komentar                                                                                                                                                                                           | label_aktual | label_prediksi |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------------|
| yakin mereka ngak bawa motor? sy kok curiga ada sewa parkir dadakan                                                                                                                                | negatif      | negatif        |
| Keren, wajib berlaku di seluruh Indonesia ini                                                                                                                                                      | positif      | positif        |
| Yg 11 km dari sekolah                                                                                                                                                                              | negatif      | negatif        |
| Gubernur jkt lagi sibuk mindahin patung dan buat patung baru,sambil ngejar ngejar penunggak pajak kendaraan bermotor??                                                                             | netral       | negatif        |
| Sangat setuju banget sama sangat cocok bgt ya.... Lbh baik satu tahun aja ya..... Bravo kang Dedi Mulyadi                                                                                          | positif      | positif        |
| dari umur berapa???                                                                                                                                                                                | netral       | netral         |
| mau jam brp GK jadi masalah yg jadi masalah pendidikan GK gratis di Malaysia swasta ja gratis 10 tahun Jokowi bikin pendidikan ketinggalan jauh dr negara Malaysia                                 | negatif      | negatif        |
| masuk jam 6.30 dah di anggap beban gua di jakarta aja masuk jam 6.00 pulang jam 16.00 ga jadi beban, karna udah biasa dan ga males malesan itu juga buat belajar di siplin dengan waktu, bagus itu | positif      | positif        |
| Kenapa orang dewasa suka terobsesi menerapkan standar dirinya pada anak <sup>2</sup> . Kenapa tidak memperhitungkan hak anak. Mentang <sup>2</sup> anak <sup>2</sup> gak punya kekuatan            | negatif      | negatif        |
| Ga ngaruh di sekolahku tetep jam 13.15                                                                                                                                                             | negatif      | negatif        |
| 6.30 normal kali                                                                                                                                                                                   | positif      | positif        |
| setujuuuuuu banget!!!!                                                                                                                                                                             | positif      | positif        |
| cirebon banyak bagt,,                                                                                                                                                                              | negatif      | negatif        |
| banjir we nu di bahas teh itu hukum alam banjir mah                                                                                                                                                | netral       | netral         |
| banyak yg protes pak karena istilah tidak berbau agama coba kalau pake kata sedekah                                                                                                                | netral       | netral         |
| mantap gubernur Jabar                                                                                                                                                                              | positif      | positif        |
| udah pajak seglanya udah di potong langsung ngapain juga                                                                                                                                           | negatif      | positif        |

Berdasarkan hasil pengujian model terhadap data komentar masyarakat, tabel di atas menampilkan perbandingan antara label aktual (hasil pelabelan berbasis lexicon) dan label prediksi (hasil klasifikasi menggunakan model LSTM).

Secara umum, sebagian besar komentar berhasil diklasifikasikan dengan benar, khususnya pada komentar dengan sentimen yang jelas dan eksplisit. Komentar bernada positif seperti *“Keren, wajib berlaku di seluruh Indonesia ini”* dan *“mantap gubernur Jabar”* berhasil diprediksi dengan tepat sebagai sentimen positif. Hal yang sama juga terjadi pada komentar negatif eksplisit, misalnya *“Yg 11 km dari sekolah”* dan *“Ga ngaruh di sekolahku tetep jam 13.15”*, yang konsisten diklasifikasikan sebagai negatif.

Komentar netral yang bersifat informatif atau berupa pertanyaan singkat, seperti *“dari umur berapa???”* dan *“banjir we nu di bahas teh itu hukum alam banjir mah”*, juga dapat dikenali dengan cukup baik oleh model.

Namun demikian, terdapat beberapa kasus salah klasifikasi, terutama pada komentar yang mengandung:

- ironi atau kritik implisit, dan
- kata positif tetapi konteksnya negatif.

Contohnya, komentar “*Gubernur jkt lagi sibuk mindahin patung dan buat patung baru...*” yang berlabel aktual netral diprediksi sebagai negatif. Hal ini menunjukkan bahwa model cenderung menangkap kata-kata bernuansa kritik sebagai sentimen negatif meskipun konteksnya masih bersifat deskriptif.

Kasus lain yang menarik adalah komentar “*udah pajak seglanya udah di potong langsung ngapain juga*” yang berlabel aktual negatif namun diprediksi positif. Kesalahan ini diduga terjadi karena adanya kata atau frasa yang tidak sepenuhnya tertangkap konteks emosinya oleh model, sehingga makna keluhan tidak terdeteksi secara optimal.

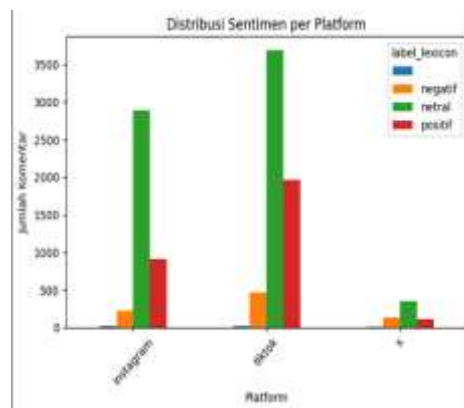
Hasil ini menunjukkan bahwa model LSTM cukup baik dalam mengklasifikasikan sentimen dengan ekspresi yang tegas, baik positif maupun negatif. Namun, model masih mengalami keterbatasan dalam memahami konteks implisit, sarkasme, dan kritik halus, yang menyebabkan sebagian kecil kesalahan prediksi. Oleh karena itu, hasil klasifikasi ini perlu diinterpretasikan secara kontekstual dan dapat ditingkatkan di masa depan melalui penyeimbangan data, penambahan kamus konteks, atau pendekatan semantik lanjutan.

## 2) Distribusi Komentar

Pada tahap interpretasi hasil, penelitian ini menyajikan informasi berupa distribusi sentimen komentar terhadap kebijakan pemerintah yang dianalisis. Distribusi sentimen ini diperoleh berdasarkan hasil proses analisis data yang telah dilakukan pada dataset komentar masyarakat yang dikumpulkan dari berbagai sumber media sosial.

### 1. Distribusi platform

Berikut Gambar Dari distribusi sentiment dengan filter per platform pada gambar 4.2



**Gambar 3.** Gambar Distribusi Sentimen Per Platform

Berikut Tabel distribusi table filter perplatform pada tabel 11

**Tabel 11.** Tabel Distribusi perplatform

| no | platform  | netral | positif | negatif |
|----|-----------|--------|---------|---------|
| 1  | instagram | 2725   | 1050    | 277     |
| 2  | tiktok    | 3600   | 1918    | 620     |
| 3  | twitter   | 351    | 126     | 127     |

Distribusi sentimen berdasarkan platform media sosial menunjukkan perbedaan pola respons pengguna terhadap topik yang dianalisis. Berdasarkan hasil pengolahan data, sentimen netral mendominasi pada seluruh platform yang diamati, yaitu Instagram, Twitter, dan TikTok.

Pada platform Instagram, mayoritas komentar pengguna bersentimen netral sebanyak 2.725 data, diikuti oleh sentimen positif sebanyak 1.050 data, dan sentimen negatif sebanyak 277 data. Hal ini menunjukkan bahwa sebagian besar pengguna Instagram memberikan tanggapan yang bersifat informatif atau tidak menunjukkan kecenderungan emosi yang kuat.

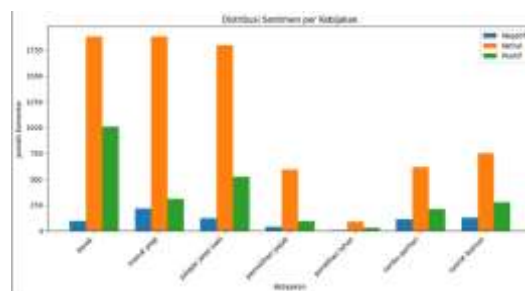
Selanjutnya, pada platform Tiktok, jumlah sentimen netral tercatat paling tinggi dibandingkan platform lainnya, yaitu sebanyak 3.600 data. Sentimen positif pada Tiktok juga cukup signifikan dengan 1.916 data, sedangkan sentimen negatif tercatat sebanyak 620 data. Kondisi ini menunjukkan bahwa Tiktok menjadi platform dengan tingkat partisipasi pengguna yang tinggi serta variasi sentimen yang lebih beragam.

Sementara itu, pada platform Twitter, jumlah komentar yang dianalisis relatif lebih sedikit dibandingkan Instagram dan Tiktok. Sentimen netral tetap mendominasi dengan 351 data, diikuti oleh sentimen positif sebanyak 126 data, dan sentimen negatif sebanyak 127 data. Meskipun jumlah data lebih kecil, proporsi sentimen negatif pada platform Twitter terlihat lebih menonjol dibandingkan dua platform lainnya.

Secara keseluruhan, hasil distribusi sentimen per platform menunjukkan bahwa sentimen netral merupakan sentimen yang paling dominan pada seluruh platform media sosial. Hal ini mengindikasikan bahwa mayoritas pengguna memberikan komentar yang bersifat objektif atau informatif, sementara sentimen positif dan negatif muncul sebagai respons terhadap pengalaman atau persepsi tertentu dari masing-masing pengguna platform.

## 2. Distribusi kebijakan

Berikut Gambar distribusi sentiment filter per kebijakan pada gambar 4.3



**Gambar 4.** Visualisasi Distribusi Sentimen Per Kebijakan

Berikut Tabel distribusi sentiment per kebijakan pada tabel 12

**Tabel 12.** Tabel Distribusi Sentimen Per Kebijakan

| Kebijakan                | Total Komentar |
|--------------------------|----------------|
| Barak                    | 2.987          |
| Masuk Pagi               | 2.408          |
| Pelajar Jalan Kaki       | 2.445          |
| Pemutihan Pajak          | 727            |
| Peralihan Lahan          | 127            |
| Seribu Per Hari          | 941            |
| Syarat Bansos            | 1.159          |
| <b>Total Keseluruhan</b> | <b>10.794</b>  |

Distribusi sentimen terhadap berbagai kebijakan menunjukkan variasi respons pengguna yang cukup berbeda-beda. Dari hasil analisis, terlihat bahwa sentimen netral mendominasi hampir di seluruh topik, menandakan bahwa sebagian besar komentar bersifat informatif atau tidak mengekspresikan emosi yang kuat.

Kebijakan Barak memperoleh 1.663 komentar netral, 1.212 komentar positif, dan 112 komentar negatif. Dominasi sentimen netral menunjukkan bahwa sebagian besar pengguna memberikan tanggapan yang bersifat informatif atau deskriptif, meskipun dukungan positif terhadap kebijakan ini juga tergolong tinggi.

Pada kebijakan Masuk Pagi (pukul 06.30), mayoritas komentar bersentimen netral sebanyak 1.721 komentar, diikuti oleh 359 komentar negatif dan 328 komentar positif. Hal ini mengindikasikan bahwa meskipun banyak pengguna bersikap netral, terdapat kecenderungan ketidakpuasan dari sebagian masyarakat terhadap kebijakan tersebut.

Kebijakan Pelajar Jalan Kaki menunjukkan distribusi sentimen berupa 1.610 komentar netral, 700 komentar positif, dan 135 komentar negatif. Hasil ini menunjukkan bahwa tanggapan netral masih mendominasi, namun dukungan positif terhadap kebijakan ini relatif cukup signifikan dibandingkan sentimen negatif.

Pada kebijakan Pemutihan Pajak, diperoleh 499 komentar netral, 159 komentar positif, dan 69 komentar negatif. Dominasi sentimen netral menunjukkan bahwa kebijakan ini sebagian besar ditanggapi secara objektif, meskipun terdapat kecenderungan respon positif dari sebagian pengguna.

Kebijakan Peralihan Lahan memiliki jumlah komentar yang relatif sedikit, namun distribusi sentimennya lebih seimbang, yaitu 52 komentar positif, 47 komentar negatif, dan 28 komentar netral. Kondisi ini menunjukkan bahwa kebijakan tersebut memicu pro dan kontra yang cukup kuat di kalangan pengguna.

Untuk kebijakan Seribu Per Hari, tercatat 539 komentar netral, 271 komentar positif, dan 131 komentar negatif. Hasil ini menunjukkan bahwa meskipun sentimen netral mendominasi, terdapat dukungan positif yang cukup besar terhadap kebijakan ini.

Terakhir, kebijakan Syarat Bansos memperoleh 616 komentar netral, 372 komentar positif, dan 171 komentar negatif. Hal ini menandakan bahwa sebagian besar pengguna memberikan tanggapan yang bersifat netral, namun terdapat proporsi sentimen positif yang cukup menonjol dibandingkan sentimen negatif.

Secara keseluruhan, hasil distribusi sentimen menunjukkan bahwa sentimen netral masih menjadi respons paling dominan di berbagai topik kebijakan. Sentimen positif muncul sebagai dukungan terhadap kebijakan tertentu, sementara sentimen negatif menandakan adanya ketidakpuasan atau kritik dari sebagian pengguna.

## **h. Analisis Hasil**

Pada bagian ini disajikan hasil dari implementasi sistem analisis sentimen terhadap kebijakan Gubernur Jawa Barat yang telah dikembangkan menggunakan algoritma LSTM (Long Short-Term Memory). Penjelasan dapat diuraikan sebagai berikut:

Sistem ini memproses komentar masyarakat yang dikumpulkan dari berbagai platform media sosial untuk menentukan sentimen setiap komentar, apakah bersifat positif, negatif, atau netral. Hasil analisis menunjukkan bahwa mayoritas komentar bersifat netral, yang mengindikasikan bahwa sebagian besar masyarakat memberikan tanggapan yang informatif atau objektif terhadap kebijakan yang dianalisis. Sentimen positif dan negatif muncul sebagai respons terhadap persepsi atau pengalaman tertentu dari masyarakat terhadap kebijakan Gubernur. Distribusi sentimen ini kemudian divisualisasikan untuk setiap topik kebijakan agar dapat memberikan gambaran yang jelas mengenai respons publik secara keseluruhan.

### **1) implementasi model**

#### **1. Model Program**

Pada bagian ini disajikan hasil implementasi sistem analisis sentimen terhadap kebijakan Gubernur Jawa Barat yang dikembangkan menggunakan model Long Short-Term Memory (LSTM) berbasis deep learning. Sistem ini bertujuan untuk mengklasifikasikan komentar masyarakat ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral.

Sistem diimplementasikan dalam bentuk aplikasi web menggunakan framework Flask, yang memungkinkan pengguna melakukan analisis sentimen baik secara manual maupun secara massal melalui unggahan file Excel. Proses analisis sentimen dilakukan melalui beberapa tahapan sebagai berikut:

a) Preprocessing teks

Komentar yang diberikan pengguna terlebih dahulu melalui tahap preprocessing untuk meningkatkan kualitas data masukan. Tahapan ini meliputi pengubahan teks menjadi huruf kecil (lowercase), penghapusan URL, mention, serta karakter non-alfabet. Selanjutnya, dilakukan penghapusan stopwords menggunakan pustaka Sastrawi, dengan pengecualian pada kata negasi seperti tidak, bukan, belum, dan kurang karena kata tersebut dapat memengaruhi makna sentimen. Setelah itu, dilakukan proses stemming untuk mengubah kata ke bentuk dasarnya.

b) Transformasi teks ke bentuk numerik

Teks hasil preprocessing kemudian diubah menjadi bentuk numerik menggunakan tokenizer yang sama dengan yang digunakan pada tahap pelatihan model. Hasil tokenisasi selanjutnya dipadatkan (padding) hingga mencapai panjang maksimum sequence yang telah ditentukan, yaitu 100 token, agar sesuai dengan input model LSTM.

c) Prediksi sentimen menggunakan model LSTM

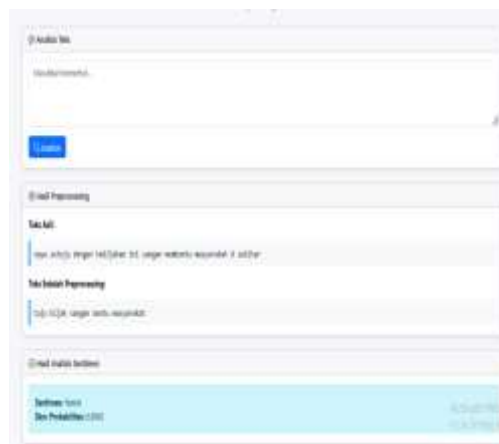
Data teks yang telah diproses kemudian dimasukkan ke dalam model LSTM untuk menghasilkan probabilitas masing-masing kelas sentimen. Kelas dengan nilai probabilitas tertinggi dipilih sebagai label sentimen akhir, yaitu positif, netral, atau negatif. Selain label sentimen, sistem juga menampilkan nilai confidence sebagai tingkat kepercayaan model terhadap hasil prediksi tersebut.

d) Fitur analisis data secara massal

Sistem menyediakan fitur unggah file Excel yang berisi kolom komentar. Setiap komentar dalam file tersebut diproses secara otomatis menggunakan tahapan preprocessing dan prediksi yang sama. Hasil analisis ditampilkan dalam bentuk tabel yang memuat teks hasil preprocessing, label sentimen, serta nilai confidence untuk setiap komentar.

e) Output sistem

Output yang dihasilkan oleh sistem berupa teks hasil preprocessing, label sentimen (positif, netral, atau negatif), serta nilai confidence hasil prediksi model. Untuk analisis massal, hasil ditampilkan dalam bentuk tabel pada halaman web sehingga memudahkan pengguna dalam melakukan evaluasi dan interpretasi hasil analisis sentiment.



**Gambar 5.** Tampilan hasil Sentimen inputan Text

[illegible]

**Gambar 6.** Tampilan Hasil Inputan Excel



**Gambar 7.** Tampilan Hasil WordCloud inputan Excel

## 2. Tampilan Website

Beberapa bagian utama dari halaman web ini antara lain:

Navbar:

Menampilkan judul aplikasi “Sentimen LSTM” dengan ikon, serta memberikan identitas visual yang konsisten di seluruh halaman.

Input Teks:

Pengguna dapat langsung memasukkan komentar di kolom textarea untuk dianalisis. Terdapat tombol “Analisis” yang akan memicu POST request ke backend Flask untuk memproses teks tersebut.

Hasil Preprocessing:

Setelah komentar diproses, halaman menampilkan teks asli dan teks hasil preprocessing, termasuk normalisasi, penghapusan stopwords, dan stemming. Bagian ini membantu pengguna memahami bagaimana sistem mempersiapkan data sebelum prediksi sentimen.

Hasil Analisis Sentimen:

Menampilkan label sentimen (Positif, Negatif, atau Netral) beserta skor probabilitas dari model LSTM. Bagian ini menggunakan komponen alert dari Bootstrap agar hasil terlihat jelas dan menarik.

#### Upload Excel:

Pengguna dapat mengunggah file Excel yang berisi banyak komentar sekaligus. Sistem akan membaca kolom komentar, memproses setiap baris, dan menampilkan hasil dalam bentuk tabel interaktif di halaman web.

#### Hasil Analisis Excel & WordCloud:

Jika file Excel diunggah dan diproses, hasilnya ditampilkan dalam tabel HTML yang responsif. Selain itu, halaman dapat menampilkan WordCloud untuk memvisualisasikan kata-kata yang paling sering muncul dari komentar yang dianalisis, memberikan insight tambahan mengenai opini publik.

Desain menggunakan Bootstrap 5 untuk tata letak responsif, ikon Bootstrap Icons untuk visualisasi interaktif, dan CSS khusus untuk memberikan sentuhan modern seperti card shadow, rounded corners, dan pre-box untuk teks preprocessed. Halaman ini memberikan pengalaman pengguna yang mudah dipahami sekaligus mendukung analisis sentimen secara real-time maupun batch.Output:



**Gambar 8.** Tampilan Halaman Utama Website Sentimen

### i. Pengujian

#### 1) Pengujian Fitur Aplikasi

Dalam pembahasan ini, penulis melakukan pengujian terhadap aplikasi analisis sentimen komentar masyarakat terhadap kebijakan Gubernur Jawa Barat berdasarkan spesifikasi sistem yang telah dibahas pada bab sebelumnya. Metode pengujian yang digunakan adalah Black Box Testing, yang berfokus pada pengujian spesifikasi fungsional sistem tanpa melihat implementasi internal.

Pengujian dilakukan pada berbagai fitur utama aplikasi, seperti input komentar secara manual, upload file Excel/CSV, proses preprocessing, prediksi sentimen menggunakan Model LSTM, serta visualisasi hasil dalam bentuk tabel dan WordCloud. Hasil pengujian menunjukkan bahwa seluruh fitur berjalan sesuai harapan dan aplikasi mampu memproses komentar secara akurat dan responsif.

**Tabel 12.** Tabel Hasil Pengujian

| No | Fitur         | Skenario Pengujian                                                | Input                                               | hasil    |
|----|---------------|-------------------------------------------------------------------|-----------------------------------------------------|----------|
| 1  | Upload Data   | Mengunggah dataset komentar masyarakat dalam format yang didukung | File Excel/CSV berisi kolom komentar                | berhasil |
| 2  | Validasi Data | Mengunggah file dengan format salah atau kolom komentar kosong    | File Excel/CSV selain kolom komentar atau tidak ada | berhasil |
| 3  | Analisis Teks | Memproses komentar yang dimasukkan secara manual                  | Teks komentar di form input                         | berhasil |

|   |                 |                                                                    |                                                           |          |
|---|-----------------|--------------------------------------------------------------------|-----------------------------------------------------------|----------|
| 4 | Analisis Batch  | Melakukan analisis sentimen untuk semua komentar dalam file Excel  | Dataset komentar                                          | berhasil |
| 5 | Prediksi Hybrid | Memastikan prediksi menggunakan kombinasi LSTM dan lexicon scoring | Komentar pendek, panjang, atau mengandung kata/frasa kuat | berhasil |

#### 4. Kesimpulan

Penelitian ini berhasil merancang dan membangun sistem berbasis web untuk melakukan analisis sentimen komentar masyarakat terhadap kebijakan Gubernur Jawa Barat dengan menggunakan metode LSTM. Sistem ini mampu memproses komentar baik secara manual maupun melalui unggahan file Excel, dan menampilkan hasil berupa label sentimen (Positif, Negatif, Netral), skor probabilitas, serta visualisasi tambahan seperti tabel hasil dan WordCloud. Dataset penelitian mencakup 10.794 komentar yang diperoleh dari tiga platform media sosial: Instagram (4.052 komentar), Tiktok (6.138 komentar), dan Twitter (604 komentar). Setelah melalui proses preprocessing dan pelabelan, data dibagi dengan rasio 80:20 untuk pelatihan dan pengujian model LSTM. Hasil pengujian menunjukkan bahwa model LSTM dengan rasio pembagian data 80:20 menghasilkan performa terbaik dengan akurasi 79,79%, precision 0,83, recall 0,83, dan F1-score 0,82.

#### Referensi

- [1] Kementerian PANRB, 'Pedoman Standar Penyusunan Survei Kepuasan Masyarakat,' 2017.
- [2] Tirta.id, 'Wagub DIY Raih Indeks Kepuasan Warga Tertinggi Sepuluh Jawa,' 28 Mei 2025.
- [3] A. Ceron, L. Curini, and S. M. Iacus, 'Using sentiment analysis to monitor electoral campaigns: Method matters—evidence from the United States and Italy,' *Social Science Computer Review*, vol. 33, no. 1, pp. 3–20, 2015.
- [4] I. Mergel, 'Social media adoption and resulting tactics in the US federal government,' *Government Information Quarterly*, vol. 30, no. 2, pp. 123–130, 2013.
- [5] S. Stieglitz and L. Dang-Xuan, 'Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior,' *Journal of Management Information Systems*, vol. 29, no. 4, pp. 217–248, 2013.
- [6] Y. Bae and H. Lee, 'Sentiment analysis of Twitter audiences: Measuring the positive or negative influence of popular twitterers,' *J. Assoc. Inf. Sci. Technol.*, vol. 63, no. 12, pp. 2521–2535, 2012.
- [7] Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), *Laporan Survei Internet APJII 2024*. Jakarta, Indonesia: APJII, 2024.
- [8] We Are Social and Meltwater, *Digital 2024: Indonesia*. London, UK: We Are Social, 2024.
- [9] Badan Pusat Statistik Provinsi Jawa Barat, *Provinsi Jawa Barat Dalam Angka 2023*. Bandung, Indonesia: BPS Jawa Barat, 2023.
- [10] Badan Pusat Statistik, *Produk Domestik Regional Bruto Menurut Provinsi 2019–2023*. Jakarta, Indonesia: BPS, 2024.
- [11] Dinas Kebudayaan dan Pariwisata Jawa Barat, *Profil Kebudayaan Provinsi Jawa Barat 2023*. Bandung, Indonesia: Disbudpar Jabar, 2023.
- [12] S. Stieglitz and L. Dang-Xuan, "Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior," *J. Manage. Inf. Syst.*, vol. 29, no. 4, pp. 217–248, 2013.
- [13] M. A. Fauzi and R. P. Utomo, "Analisis sentimen pada media sosial menggunakan metode Naïve Bayes dan Support Vector Machine," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 4, pp. 401–408, 2019.
- [14] R. Kurniawan and D. K. Dewi, "Analisis sentimen terhadap kebijakan pemerintah pada Twitter menggunakan metode SVM," *J. Edukasi dan Penelitian Informatika (JEPIN)*, vol. 5, no. 2, pp. 217–224, 2019.
- [15] M. C. Farsiah, "Analisis Sentimen Menggunakan Arsitektur LSTM Terhadap Fenomena Citayam Fashion Week," 2023.
- [16] A. Kustiwa, A. F. Akbar, and A. Pinandito, "Studi Perbandingan CNN-LSTM dan LSTM dalam Mendeteksi Emosi pada Data Teks," 2022.
- [17] Y. A. Pradana, I. Cholissodin, and D. Kurnianingtyas, "Analisis Sentimen Pemindahan Ibu Kota Indonesia menggunakan LSTM dan Word2Vec," 2022.
- [18] M. N. I. Sina and E. B. Setiawan, "Stock Price Correlation Analysis with Twitter Sentiment Using CNN-LSTM," 2023.
- [19] S. F. Handayani, R. W. Pratiwi, Dairoh, and D. I. Af'idah, "Analisis Sentimen pada Ulasan Twitter dengan LSTM," 2021.
- [20] M. F. Naufal and S. F. Kusuma, "Analisis Sentimen Twitter terhadap Kebijakan PPKM," 2022.
- [21] D. A. Firdlous and R. Andrian, "Sentiment Analysis Public Twitter on 2024 Election using LSTM," 2024.
- [22] P. H. Prastyo, M. A. Ridho, and S. A. Wibowo, "Sentiment Analysis on COVID-19 Using SVM," 2021.
- [23] C.-H. Lin and U. Nuha, "Sentiment Analysis of Indonesian Datasets Based on Hybrid Deep Learning Strategy," 2023.
- [24] P. R. A. Savitri, I. M. A. D. Suarjaya, and W. O. Vihikan, "Sentiment Analysis on Korean Culture in Indonesia," 2022.
- [25] A. Hermida, "Twittering the news: The emergence of ambient journalism," *Journalism Practice*, vol. 4, no. 3, pp. 297–308, 2010.
- [26] A. Hu, L. Zhao, and S. Huang, "Understanding User Engagement on Instagram: A Case Study of Social Movements," *International Journal of Communication*, vol. 14, pp. 4534–4556, 2020.
- [27] J. A. Chewning and A. M. Doerfel, "Twitter and health policy: Using a network approach to explore information flow about the Affordable Care Act," *Communication Studies*, vol. 64, no. 5, pp. 451–468, 2013.
- [28] D. M. Boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship," *J. Comput.-Mediat. Commun.*, vol. 13, no. 1, pp. 210–230, 2007.
- [29] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Business Horizons*, vol. 53, no. 1, pp. 59–68, Jan. 2010.

- [30] B. Liu, "Sentiment Analysis and Opinion Mining". San Rafael, CA: Morgan & Claypool, 2012.
- [31] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *\*Neural Computation\**, vol. 9, no. 8, pp. 1735–1780, 1997.
- [33] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural computation*, vol. 12, no. 10, pp. 2451-2471, 2000.
- [34] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based LSTM for aspect-level sentiment classification," in *Proceedings of the 2016 conference on empirical methods in natural language processing*, 2016, pp. 606-615.
- [35] X. Wang, W. Jiang, and Z. Luo, "Combination of convolutional and recurrent neural network for sentiment analysis of short texts," in *Proceedings of COLING 2016, the 26th international conference on computational linguistics: technical papers*, 2016, pp. 2428-2437.
- [36] D. Jurafsky and J. H. Martin, *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition*, 3rd ed. London: Pearson Education, 2020.
- [37] C. D. Manning and H. Schütze, *Foundations of statistical natural language processing*. Cambridge: MIT Press, 2019.
- [38] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*. Sebastopol: O'Reilly Media, 2019.
- [39] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
- [40] J. S. Dryzek, "Deliberative democracy and beyond: liberals, critics, contestations," *Political Studies*, vol. 50, no. 1, pp. 1-18, 2002.
- [41] S. Grimmelikhuijsen and A. Meijer, "Effects of transparency on the perceived trustworthiness of a government organization: Evidence from an online experiment," *Journal of Public Administration Research and Theory*, vol. 24, no. 1, pp. 137-157, 2014.
- [42] Kurniawati, T. R., "Analisis Big Data dalam Pengambilan Kebijakan Publik di Indonesia," *Jurnal Ilmu Administrasi Negara*, vol. 10, no. 2, pp. 35–46, 2023.
- [43] J. Bhandari and S. Bimo, "TikTok and the Public Sphere: Short Video Platforms as New Spaces for Civic Participation," *New Media & Society*, vol. 25, no. 1, pp. 23–41, 2023.