

Analisis Sentimen Mobil Listrik Pada X (*Twitter*) Menggunakan Metode *Long Short-Term Memory* (LSTM)

Muhammad Jilan Hilmi
Fakultas Informatika
Universitas Telkom
Bandung

jilanhilmi@student.telkomuniversity.ac.id

Hani Nurrahmi
Fakultas Informatika
Universitas Telkom
Bandung

haninurrahmi@telkomuniversity.ac.id

Hasmawati
Fakultas Informatika
Universitas Telkom
Bandung

hasmawati@telkomuniversity.ac.id

Pertumbuhan kendaraan listrik di Indonesia menghadapi tantangan, salah satunya adalah rendahnya adopsi oleh masyarakat. Untuk memahami persepsi publik, dilakukan analisis sentimen berdasarkan opini masyarakat di media sosial X (*Twitter*) dengan kata kunci mobil listrik. Metode yang digunakan pada penelitian ini adalah Long Short-Term Memory (LSTM), salah satu algoritma deep learning yang efektif untuk mengklasifikasikan data berbasis teks. Data diambil dari *Twitter* menggunakan kata kunci "Mobil Listrik" dalam rentang Januari 2023 hingga April 2024, dengan total 10,283 tweet. Setiap tweet divalidasi oleh lima responden dan dikategorikan ke dalam sentimen negatif, netral, atau positif berdasarkan mayoritas suara. Proses pengujian dilakukan menggunakan tiga skenario split dataset (70:30, 80:20, dan 90:10) dan dievaluasi menggunakan metrik akurasi, precision, recall, dan f1-score. Hasil pengujian menunjukkan bahwa model LSTM memiliki akurasi pengujian sebesar 55,30% dan akurasi validasi sebesar 57,13%. Model hanya mampu mengenali sentimen netral dengan baik (f1-score: 71%), namun gagal mengklasifikasikan sentimen positif dan negatif. Hal ini menunjukkan adanya ketidakseimbangan kelas, sehingga dibutuhkan upaya perbaikan seperti penyeimbangan data dan optimasi model agar hasil klasifikasi lebih merata.

Kata kunci: analisis sentimen, mobil listrik, LSTM, *Twitter*, klasifikasi.

The adoption of electric vehicles in Indonesia remains relatively low despite their increasing popularity. To understand public perception, a sentiment analysis was conducted based on public opinion on social media X (*Twitter*) with the electric car keyword. The method used is Long Short-Term Memory (LSTM), a deep learning algorithm effective for text-based classification tasks. Data was collected using the keyword "Mobil Listrik" (Electric Car) from January 2023 to April 2024, resulting in a total of 10,283 tweets. Each tweet was manually labeled by five respondents and categorized into negative, neutral, or positive sentiment based on majority voting. The model was tested under three dataset split scenarios (70:30, 80:20, and 90:10), and evaluated using accuracy, precision, recall, and f1-score metrics. The results show that the LSTM model achieved a test accuracy of 55.30% and a validation accuracy of 57.13%. The model was only able to correctly classify neutral sentiment (f1-score: 71%), while failing to identify both positive and negative sentiments. This indicates a class imbalance issue, suggesting the need for improvements such as data balancing and model optimization to achieve more evenly distributed classification results.

Keywords: sentiment analysis, electric vehicle, LSTM, *Twitter*, classification

1. Pendahuluan

Latar Belakang

Pesatnya perkembangan teknologi di era digital saat ini, berbagai sektor kehidupan mengalami perubahan yang cukup signifikan, termasuk sektor otomotif. Salah satu inovasi yang semakin mendapat perhatian adalah mobil listrik, yakni kendaraan yang digerakkan oleh listrik dan menggunakan baterai sebagai sumber energinya. Mobil listrik memiliki sejumlah keunggulan, seperti efisiensi dalam penggunaan energi serta kontribusi positif terhadap lingkungan. Di Indonesia, adopsi mobil listrik masih berada pada tingkat yang relatif rendah. Analisis sentimen digunakan untuk menggali persepsi masyarakat terhadap mobil listrik sebagai salah satu upaya untuk mendukung pengembangan dan peningkatan adopsi kendaraan ramah lingkungan di dalam negeri[1]. Mobil listrik pertama yang bersifat praktis telah dikembangkan sejak tahun 1880-an. Kendaraan ini menggunakan listrik sebagai sumber utama energinya. Berbeda dengan kendaraan konvensional, mobil listrik hanya mengandalkan baterai sebagai sumber tenaga, sehingga tidak menghasilkan polusi udara dan mampu menghasilkan daya sebesar 3000 watt dengan tegangan 72 volt. Mobil listrik memiliki berbagai keunggulan dibandingkan kendaraan berbahan bakar konvensional. Salah satu keunggulan utamanya adalah tidak menghasilkan emisi gas buang, sehingga lebih ramah terhadap lingkungan. Selain itu, karena tidak menggunakan bahan bakar fosil, mobil listrik juga berperan dalam menurunkan emisi gas rumah kaca[2]. Produksi massal dan pemasaran komersial mobil listrik pertama kali dilakukan oleh General Motors. Kendaraan ini dirancang untuk penggunaan terbatas, dengan daya jelajah yang relatif pendek bila dibandingkan dengan kendaraan konvensional berbahan bakar fosil[3].

Respon masyarakat Indonesia terhadap mobil listrik merupakan salah satu indikator penting yang digunakan peneliti dalam menganalisis sentimen publik melalui media sosial *Twitter*[4]. Sebagai media sosial, *Twitter* digunakan oleh masyarakat untuk menyampaikan opini dan ekspresi melalui unggahan singkat yang disebut tweet [5]. Menurut laporan Statista pada bulan Juli 2021, Indonesia menduduki peringkat keenam di dunia dalam jumlah pengguna *Twitter*, dengan total sekitar 15,7 juta pengguna[6]. Popularitas *Twitter* di kalangan developer didorong oleh kemudahan dalam mengakses data yang dibutuhkan. Informasi di *twitter* tersebar melalui retweet dan like, di mana peningkatan jumlah interaksi tersebut berkontribusi pada perluasan jangkauan informasi. [7]

Analisis sentimen, atau *opinion mining*, merupakan cabang ilmu komputasi yang mempelajari opini, sentimen, dan emosi individu yang diekspresikan melalui teks terhadap suatu entitas beserta atributnya. Melalui analisis ini, opini seseorang dapat diklasifikasikan ke dalam kategori sentimen positif, negatif, atau netral terhadap permasalahan yang dibahas[8].

Dalam penelitian ini, penulis melakukan penelitian menganalisis berbagai komentar masyarakat Indonesia terkait adanya kendaraan listrik di Indonesia, dengan tujuan mengklasifikasikan opini publik terhadap isu kendaraan listrik menjadi kategori sentimen positif, negatif, dan netral. Metode yang digunakan pada penelitian ini adalah *Long Short-Term Memory* (LSTM) karena metode LSTM termasuk algoritma *deep learning* yang cocok untuk mengklasifikasikan data bentuk *text*. Penelitian ini bertujuan untuk mengklasifikasikan data *tweet* ke dalam 3 kategori, yaitu positif, negatif, dan netral.

Topik dan Batasannya

Berdasarkan pada topik penelitian yang sudah diteliti oleh penulis, maka dapat ditentukan topik masalahnya yaitu Analisis Sentimen Mobil Listrik Pada X (Twitter) menggunakan Metode *Long Short-Term Memory* (LSTM). Batasan dari penelitian ini adalah data yang diambil dan digunakan dengan kata kunci "Mobil Listrik" dalam rentang waktu Januari 2023 – April 2024.

Tujuan

Tujuan penelitian ini adalah membangun sistem Analisis Sentimen Mobil Listrik Pada X (Twitter) menggunakan Metode *Long Short-Term Memory* (LSTM) untuk mengetahui performansi serta akurasi dari metode tersebut.

Organisasi Tulisan

Struktur penulisan dalam penelitian ini disusun ke dalam beberapa bab. Bab pertama berisi uraian mengenai permasalahan yang diangkat, meliputi latar belakang, rumusan masalah, serta tujuan penelitian. Bab kedua menyajikan landasan teori yang digunakan sebagai dasar dalam menyelesaikan permasalahan, termasuk hasil penelitian terdahulu dan referensi ilmiah lainnya yang relevan. Bab ketiga menjelaskan proses perancangan sistem berdasarkan informasi yang dikaji pada bab sebelumnya. Bab keempat memaparkan hasil implementasi sistem beserta evaluasinya menggunakan beberapa *matrix* pengukuran. Terakhir, bab kelima berisi kesimpulan yang telah dilakukan penulis serta saran untuk pengembangan lebih lanjut.

2. Studi Terkait

Penulis menggunakan beberapa referensi pada penelitian yang dilakukan. Penyelidikan terkait dengan penelitian ini antara lain: Pada tahun 2022, Romzi dan Atastina melakukan sebuah studi dengan judul "Analisis Sentimen Judul Berita Ekonomi Terhadap Indeks Harga Saham Gabungan Menggunakan Metode *Long Short-Term Memory*", Dalam penelitian tersebut Romzi bertujuan untuk memprediksi pergerakan Indeks Harga Saham Gabungan (IHSG) dengan menggunakan *text to sequence* untuk merubah kata menjadi *integer*, akurasi yang dihasilkan penelitian ini adalah 49%[9].

Pada tahun 2022, Arifin dan Jondri melakukan sebuah studi dengan judul "Prediksi *Retweet* Menggunakan Fitur *User Based* dan *Content Based* dengan Metode Klasifikasi ANN-GA", Dalam penelitian tersebut Arifin bertujuan untuk memprediksi suatu *tweet* yang akan di *retweet* oleh *user* berdasarkan fitur

user-based dan *content-based*, dan mendapatkan akurasi sebesar 90%[7].

Pada tahun 2023, Fadhillah, Pramesti dan Fa'rifah melakukan penelitian dengan judul "Analisis Sentimen Masyarakat Terhadap Penggunaan Kendaraan Listrik pada Media Sosial Twitter Menggunakan Metode Algoritma *Naive Bayes*". Dalam penelitian tersebut Fadhillah bertujuan untuk mengklasifikasikan *sentimen level* dan mengimplementasikan juga metode *imbalance handling* SMOTE, dan RUS. Penelitian ini menghasilkan akurasi 94%[10].

2.1 Analisis Sentimen

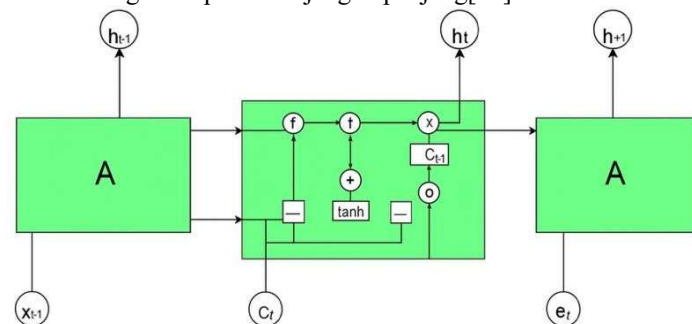
Analisis sentimen mengklasifikasikan sentimen berdasarkan polaritas teks dalam suatu frasa, sehingga dapat ditentukan apakah sentimen tersebut bersifat positif, negatif, atau netral[2]. Analisis sentimen adalah studi yang mempelajari opini, emosi, dan sikap terhadap suatu entitas seperti produk, layanan, individu, atau peristiwa. Analisis sentimen mencakup berbagai istilah seperti *opinion mining*, *sentiment extraction*, dan *emotion analysis* yang berfokus pada penggalian informasi subjektif dari teks[1].

2.2 X (Twitter)

Twitter, yang kini dikenal sebagai X, merupakan platform media sosial dan microblogging yang didirikan pada tahun 2006. Platform ini memungkinkan pengguna untuk mengirim dan membaca pesan singkat hingga 280 karakter, atau yang dikenal sebagai "*tweet*". 500 juta *tweet* per hari. *Twitter* menjadi salah satu sumber data yang banyak dimanfaatkan dalam penelitian analisis sentimen. Selain itu, *platform* ini memungkinkan interaksi antar pengguna meskipun tidak saling mengikuti[11].

2.3 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) adalah arsitektur lanjutan dari *Recurrent Neural Network* (RNN) yang dirancang khusus untuk menangani dan memproses data dalam bentuk urutan *sequential data*. LSTM hadir sebagai solusi atas permasalahan *vanishing* dan *exploding gradient* yang sering terjadi pada RNN saat menangani dependensi jangka panjang[12].



GAMBAR 1
Long Short-Term Memory (LSTM)

2.4 Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi kinerja suatu model prediksi dengan mempertimbangkan berbagai skenario pengukuran. Dalam konteks ini, sejumlah metrik digunakan untuk menilai performa model dalam memprediksi *retweet*. Evaluasi dilakukan melalui *confusion matrix* yang memuat hasil klasifikasi berupa *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [7].

TABEL 1
Confusion Matrix

Actual Class	Prediction Class	
	True	False
True	TP	TN
False	FP	FN

- *Accuracy (A)* adalah rasio dari jumlah data yang diprediksi benar, baik positif maupun negatif oleh sistem dibandingkan dengan total jumlah data yang ada pada sistem[9].

$$A = \frac{TP + TF}{TP + TN + FP + FN} (1)$$

- *Precision (P)* adalah rasio jumlah data yang diprediksi benar positif oleh system dibandingkan dengan total hasil data yang diprediksi benar pada sistem[9].

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

- *Recall (R)* adalah rasio jumlah data yang diprediksi benar positif oleh sistem dibandingkan dengan total data yang benar positif pada sistem[9].

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

- *F1-score (F1)* adalah jumlah rata-rata dari nilai *precision* dan *recall*[9].

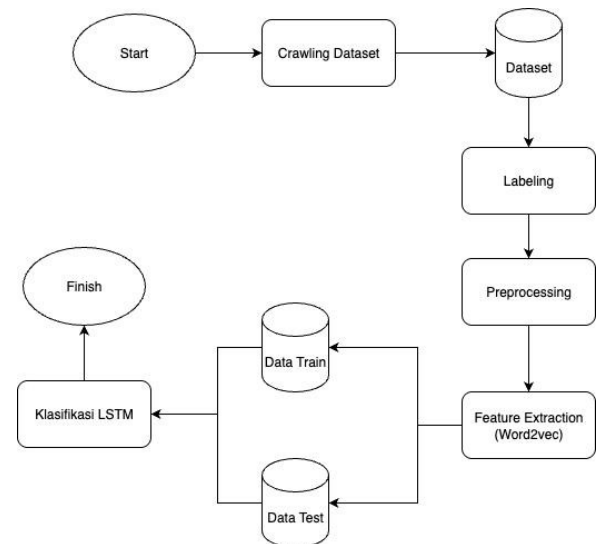
$$F1 = 2 \times \frac{(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})} (4)$$

2.5 Word2vec

Word2Vec merupakan teknik dalam pembelajaran mesin yang merepresentasikan kata ke dalam bentuk vektor numerik, dengan tujuan menangkap makna semantik dari kata dalam suatu teks. Metode ini memiliki dua arsitektur utama, yaitu *Continuous Bag-of-Words (CBOW)* dan *Skip-Gram*, yang mempelajari realasi antar kata pada korpus dan mengonversinya menjadi representasi vektor. Vektor yang dihasilkan dapat dimanfaatkan sebagai fitur dalam berbagai aplikasi pemrosesan bahasa alami (NLP) dan *machine learning*[13].

3. Sistem Yang Dibangun

Dalam penelitian ini, metode yang digunakan adalah *Long Short-Term Memory* dan mengevaluasi dengan nilai akurasi. Data yang diperoleh dari X (*Twitter*) sebanyak 10.283 data. Gambar 2 merupakan rancangan sistem yang telah dibangun. Tahap pertama dimulai dari *crawling dataset*. *Crawling* dilakukan menggunakan X (*Twitter*) *developer* dan mengubahnya menjadi *dataset* yang siap diolah. Tahap kedua adalah *preprocessing data*. Pada tahap ini dilakukan *case folding*, pembersihan data (URL, *mention*, emoji, digit, tanda baca), *tokenizing*, normalisasi dan *stemming/lemmatization*. Tahap ketiga dilakukan *feature extraction* dengan *split dataset*, dataset diubah menjadi *data train* dan *data test*. Tahap terakhir dilakukan klasifikasi menggunakan LSTM.



GAMBAR 2
Sistem yang Dibangun

3.1

Crawling Data

Crawling data yang penulis lakukan adalah menggunakan *library python* dengan Teknik *scraping* menggunakan *snsrape*. Hasil yang didapatkan 10.283 data *tweet* bahasa Indonesia dengan kata kunci “Mobil Listrik” dengan rentang waktu Januari 2023 – April 2024.

3.2

Labeling Data

Pada proses ini dataset diberikan ke 5 responden, dan responden melakukan labeling sentimen level. Responden yang dipilih memiliki latar belakang pendidikan S1 informatika dan S2 Sistem Informasi. Responden memiliki latar belakang pekerjaan sebagai Data Scientist, Data Analyst, dan Auditor. Masing - masing responden mengisi sentimen level sebanyak 10.283 data yang sudah disiapkan.

3.3

Pre-Processing

Pada tahap ini penulis melakukan *preprocessing* data dengan menggunakan *text Case Folding*, pembersihan data dan pengecekan data duplikasi dilakukan untuk memastikan tidak adanya duplikasi data, dan melakukan *tokenization*, *normalization*, *stemming*, *stopwords removal*, *Text to sequence*.

4.

Evaluasi

Penelitian ini menggunakan *dataset* yang sudah dikumpulkan berjumlah 10.283 data. Evaluasi dilakukan untuk mengetahui dalam skenario pengujian *split dataset* apakah menghasilkan nilai performansi tinggi.

4.1 Validasi Sentimen

TABEL 2
Validasi Sentimen

<i>Tweet</i>	R ₁	R ₂	R ₃	R ₄	R ₅	<i>Final</i> sentimen
Biar horangkayah pake mobil listrik subsidi horang miskin bayar subsidi mereka dengan udunan naik kereta	-1	0	-1	1	-1	-1
Bulan Depan Suzuki Luncurkan 3 Produk Baru Salah Satunya Mobil Listrik	0	-1	0	1	1	0
Kenangan baik di 2022 akan disambut dengan kebijakan naiknya tarif KRL sementara pajak mobil listrik ditekan seminim mungkin. Lu kira orang kaya mau naik kereta? Naik Tesla dkk lah	-1	1	-1	1	1	1
Kendaraan listrik akan mulai naik daun tentunya dengan fasilitas self-driving.	0	0	0	1	1	0
Mobil Listrik Kloningan Daihatsu Ini Siap Masuk Pasar Indonesia: EPMB menerima hak eksklusif dalam merakit dan mendistribusi kendaraan listrik merek Lingbox di pasar Malaysia dan Indonesia. Mobil Lingbox mini EV segera masuk ke tanah air?	0	-1	0	1	1	0

Keterangan: R₁: Responden 1, R₂: Responden 2, R₃: Responden 3, R₄: Responden 4, R₅: Responden 5.

Pada tahap ini peneliti melakukan validasi sentimen yang dilakukan oleh 5 responden, proses validasi sentimen dilakukan dengan membandingkan hasil klasifikasi sentimen dari beberapa responden. Berdasarkan **mayoritas** sentimen, mayoritas *sentimen level* akan dijadikan *final sentimen*. Jika terdapat nilai *sentimen level* dengan jumlah yang sama maka akan dimasukan ke *sentimen level* 0, yaitu sentimen netral.

4.2 Hasil Validasi Sentimen

Pada tahap ini peneliti melakukan penjumlahan validasi sentimen yang dilakukan oleh 5 responden, *final sentimen* adalah hasil sentimen terbanyak dalam satu *tweet*, atau netral jika satu *tweet* menghasilkan jumlah sentimen dengan sama. Tabel 3 menunjukkan jumlah *sentimen level* yang didapat, 0 (Netral) mendapat jumlah 5795, 1 (positif) mendapat jumlah 3263 sentimen, dan -1 (negatif) mendapat 1225 sentimen.

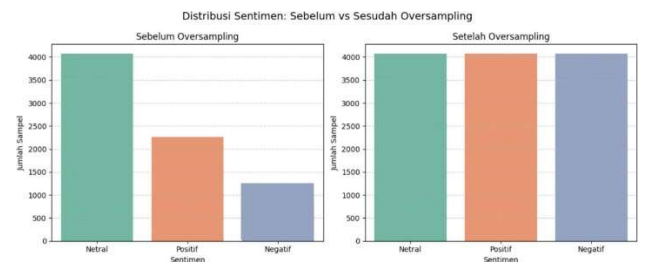
TABEL 3
Hasil Validasi Sentimen

<i>final manual sentiment</i>	<i>count</i>
0	5795
1	3263
-1	1225

4.3 Oversampling

Pada Gambar 3, grafik sebelah kiri, terlihat bahwa *total* data sentimen netral jauh lebih banyak dibandingkan dengan sentimen positif dan negatif, dengan jumlah sampel masing-masing sekitar 5795, 3263, dan 1225. Ketidakseimbangan ini berpotensi menyebabkan model pembelajaran mesin menjadi bias terhadap kelas yang dominan, sehingga performa dalam mengklasifikasikan kelas minoritas menjadi kurang optimal.

Mengatasi permasalahan tersebut, dilakukan proses *oversampling* terhadap kelas minoritas, agar jumlah data pada ketiga kategori sentimen menjadi seimbang. Hasil dari proses ini ditunjukkan pada grafik sebelah kanan, di mana ketiga kelas sentimen netral, positif, dan negatif memiliki jumlah sampel yang relatif sama, yaitu sekitar 5795.



GAMBAR 3
Oversampling

4.4 Hasil Pengujian

Skenario Split Dataset 70:30 Sebelum Oversampling

Pada tahap awal pengujian, dataset dibagi menjadi 70% untuk *data train* dan 30% untuk *data test*. Tujuan pengujian ini adalah untuk mengevaluasi performa model berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Model diuji sebanyak 10 kali untuk mencari kombinasi parameter terbaik yang menghasilkan performa optimal. Metode yang digunakan dalam proses ini adalah *Long Short-Term Memory* (LSTM).

TABEL 4
Hasil Split Dataset

Sentimen Level	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
-1	56%	0%	0%	0%
0	56%	56%	100%	72%
1	56%	0%	0%	0%

Pengujian model menggunakan metode *Long Short-Term Memory* (LSTM) sebelum menggunakan *oversampling* SMOTE dengan pembagian data 70% *training* dan 30% untuk *testing* menghasilkan akurasi validasi yang stabil. Selama *train* selama 10 *epoch*, model menunjukkan peningkatan ringan pada akurasi training dari 0.5483 hingga 0.5800, namun akurasi validasi (*val accuracy*) cenderung stagnan di sekitar 0.5750. Berdasarkan hasil evaluasi klasifikasi, model hanya mampu mengenali kelas 0 (netral) dengan baik, yaitu *precision* sebesar 56%, *recall* 100%, dan *f1-score* 72%. Sebaliknya, model gagal mengenali kelas -1 (negatif) dan 1 (positif) dengan *f1-score* masing-masing sebesar 0.00%. Nilai *macro average f1-score* sebesar 24%.

Skenario Split Dataset 70:30 Setelah Oversampling

Pada tahap awal pengujian, dataset dibagi menjadi 70% untuk *data train* dan 30% untuk *data test* menggunakan *oversampling* SMOTE. Tujuan pengujian ini adalah untuk mengevaluasi performa model berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Model diuji sebanyak 10 kali untuk mencari kombinasi parameter terbaik yang menghasilkan performa optimal. Metode yang digunakan dalam proses ini adalah *Long Short-Term Memory* (LSTM).

TABEL 5

Hasil *Split Dataset Oversampling*

Senti men Level	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
-1	56%	0%	0%	0%
0	56%	56%	100%	72%
1	56%	0%	0%	0%

Setelah dilakukan *oversampling* menggunakan SMOTE untuk menyeimbangkan distribusi kelas, model dilatih kembali untuk melakukan klasifikasi sentimen. Data dibagi menjadi 70% untuk *train* dan 30% untuk *test*. Setelah pelatihan selesai, hasil evaluasi model terhadap data uji menunjukkan akurasi pengujian sebesar 55,75% dengan nilai *loss* sebesar 43.2462. Namun demikian, jika dilihat dari laporan klasifikasi, kinerja model tetap hanya baik pada kelas 0 (netral), dengan persentase nilai *precision* sebesar 56%, *recall* 100%, dan *f1-score* sebesar 72%.

Skenario *Split Dataset 80:20* sebelum *Oversampling*

Pada pengujian kedua, dilakukan percobaan pembagian pada *dataset*, pembagian pertama 80% data *training* dan 20% data *testing*. Pengujian dilakukan untuk mengetahui hasil nilai *Accuracy*, *Precision*, *Recall* dan *F1-Score* pada *dataset*. Pengujian dilakukan sebanyak 10 kali untuk percobaan kombinasi terbaik sehingga menghasilkan nilai kombinasi yang maksimal. Metode yang digunakan pada pengujian ini adalah *Long Short-Term Memory* (LSTM).

TABEL 6
Hasil *Split Dataset*

Sentimen Level	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
-1	55%	0%	0%	0%
0	55%	55%	100%	71%
1	55%	0%	0%	0%

Pengujian model juga dilakukan dengan menggunakan skenario pembagian data sebesar 80% untuk data *train* dan 20% untuk data *test*, tanpa dilakukan proses *oversampling* (sebelum SMOTE). Model LSTM dilatih selama 10 *epoch*. Selama proses pelatihan, model menunjukkan akurasi validasi yang cukup stabil dengan rata-rata di sekitar 55,13%, dan nilai *val loss* berkisar antara 0.5013 hingga 0.5154, *test loss* 0.5046. Menunjukkan bahwa model relatif stabil dalam proses pelatihan, meskipun peningkatan kinerja tidak signifikan.

Skenario *Split Dataset 80:20* setelah *Oversampling*

Pada pengujian ini, data dibagi dengan perbandingan 80% untuk *train* dan 20% untuk *test*, kemudian dilakukan proses *oversampling* menggunakan SMOTE untuk mengatasi ketidakseimbangan jumlah data antar kelas. Dilakukan pelatihan model selama 10 *epoch*.

TABEL 7

Hasil *Split Dataset Oversampling*

Sentimen Level	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
-1	55%	0%	0%	0%
0	55%	55%	100%	71%
1	55%	0%	0%	0%

Pengujian model dengan skenario pembagian data 80% *train* dan 20% *test* setelah *oversampling* (SMOTE) menggunakan LSTM

selama 10 *epoch* menunjukkan akurasi validasi stabil di sekitar 56,87%, dengan *val_loss* antara 0.5013 hingga 0.5154. Evaluasi terhadap data *train* menghasilkan *test accuracy* sebesar 55,13% dan *test loss* 0.5046. Model hanya mampu mengenali kelas netral (kelas 0) dengan *precision* 55%, *recall* 100%, dan *f1-score* 71%.

Skenario *Split Dataset 90:10* sebelum *Oversampling*

Pada pengujian akhir, data dibagi dengan proporsi 90% untuk *train* dan 10% untuk *test*. Pengujian ini bertujuan untuk mengukur kinerja model menggunakan metrik *akurasi*, *presisi*, *recall*, dan *F1-score*. Proses pelatihan model LSTM dilakukan sebanyak 10 kali guna mengevaluasi berbagai kombinasi parameter dan memperoleh hasil terbaik.

TABEL 8
Hasil *Split Dataset*

Sentimen Level	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
-1	55%	0%	0%	0%
0	55%	55%	100%	71%
1	55%	0%	0%	0%

Pengujian model dilakukan menggunakan skenario pembagian data 90% untuk data latih dan 10% untuk data uji tanpa proses *oversampling* (SMOTE). Model yang digunakan adalah *Long Short-Term Memory* (LSTM) dan dilatih selama 10 *epoch*. Hasil pelatihan menunjukkan akurasi validasi yang konsisten sebesar 57,13%, dengan nilai *val_loss* yang relatif stabil pada rentang 0.4931 hingga 0.5001. Setelah pelatihan selesai, model dievaluasi pada data uji dan memperoleh *test accuracy* sebesar 55,30% dan *test loss* sebesar 0.5159. Evaluasi klasifikasi menunjukkan bahwa model hanya mampu mengklasifikasikan kelas netral (kelas 0) dengan baik, yaitu dengan *precision* sebesar 55%, *recall* 100%, dan *f1-score* sebesar 71%.

Skenario *Split Dataset 90:10* setelah *Oversampling*

Pada pengujian akhir, data dibagi dengan proporsi 90% untuk *training* dan 10% untuk *testing*, dengan menerapkan teknik *oversampling* menggunakan metode SMOTE untuk menangani ketidakseimbangan kelas. Pengujian ini bertujuan untuk mengukur kinerja model menggunakan metrik *akurasi*, *presisi*, *recall*, dan *F1-score*. Proses pelatihan model LSTM dilakukan sebanyak 10 kali guna mengevaluasi berbagai kombinasi parameter dan memperoleh hasil terbaik.

TABEL 9

Hasil *Split Dataset Oversampling*

Sentimen Level	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
-1	55%	0%	0%	0%
0	55%	55%	100%	71%
1	55%	0%	0%	0%

Setelah menerapkan teknik *oversampling* menggunakan metode SMOTE, data dibagi dengan proporsi 90% untuk data latih dan 10% untuk data uji. Model *Long Short-Term Memory* (LSTM) dilatih selama 10 *epoch* dan menunjukkan akurasi validasi yang cukup konsisten, yaitu sebesar 57,13%, dengan nilai *val_loss* yang berada pada rentang 0.4931 hingga 0.5001. Evaluasi terhadap data uji menunjukkan bahwa model memperoleh akurasi sebesar 55,30% dengan nilai *test loss* sebesar 0.5159.

Berdasarkan hasil klasifikasi, model hanya dapat mengenali kelas netral (kelas 0) dengan baik, ditandai oleh nilai *precision* sebesar 55%, *recall* 100%, dan *f1-score* sebesar 71%.

4.5 Analisis Pengujian

Hasil implementasi dan analisis kinerja model LSTM dalam melakukan klasifikasi sentimen terhadap *data tweet* mengenai mobil listrik. Evaluasi dilakukan dengan tiga skenario pembagian data, yaitu 70:30, 80:20, dan 90:10, baik sebelum maupun sesudah dilakukan proses *oversampling* menggunakan metode SMOTE. Metrik evaluasi yang digunakan meliputi akurasi, *precision*, *recall*, dan *f1-score* untuk kelas sentimen positif, netral, dan negatif. Hasil pengujian sebelum dilakukan *oversampling* menunjukkan model cenderung bias terhadap kelas mayoritas, yaitu sentimen netral. Pada semua skenario pembagian data 70:30, 80:20, dan 90:10, akurasi model berada di kisaran 55%–56%, dan hanya kelas netral (label 0) yang mampu dikenali dengan baik, dengan *f1-score* berkisar antara 71%–72%. Sebaliknya, model sama sekali tidak mampu mengenali kelas sentimen negatif (label -1) dan positif (label 1), yang ditunjukkan oleh nilai *precision*, *recall*, dan *f1-score* sebesar 0%. Hal tersebut menunjukkan bahwa model mengalami masalah *imbalance class* yang cukup serius, karena dominasi data netral sangat memengaruhi proses pembelajaran model. Dilakukan proses *oversampling* menggunakan metode SMOTE. Setelah dilakukan *oversampling* dan distribusi data menjadi seimbang, model kembali diuji dengan skenario split yang sama. Akan tetapi, hasil evaluasi menunjukkan bahwa kinerja model tidak mengalami peningkatan signifikan. Akurasi *model* tetap berada pada kisaran 55%–56%, dan *model* masih hanya mampu mengenali kelas netral, sementara sentimen negatif dan positif tetap tidak berhasil diklasifikasikan dengan baik. Meskipun distribusi data telah diperbaiki, hal ini menunjukkan bahwa *oversampling* saja belum cukup untuk meningkatkan kemampuan model dalam mengenali pola dari kelas minoritas.

TABEL 10

Perbandingan Validasi Sentimen Sebelum dan Sesudah Klasifikasi LSTM

Full_Text	Final_Manual_Sentiment	Predicted_Sentiment
Produsen mobil listrik Great Wall Motors (GWM) memamerkan salah satu mobilnya yakni GWM ORA 3 di GIIAS 2023 ICE BSD. Mobil ini memiliki bentuk yang unik dan menarik perhatian.	0	0
efisiensi ketika nantinya Electrizen ingin beralih ke kendaraan listrik. Selain itu Electrizen juga bisa lebih up to date mendapat informasi terkini terkait motor dan mobil listrik dan menemukan titik-titik SPKBLU di sekitar.	0	0
BYD Resmi Masuk Indonesia Pekan	0	0

Depan Mobil Listrik Terlaris Dunia		
Sebelumnya kalian udah tau belum kalau lagi ada BRImo FSTVL bisa kasih kamu banyak benefit. Salah satunya yaitu kamu bisa menangin mobil listrik Mini Electric dan MG4 EV. Penasaran gimana cara dapetnya?	0	0
China sudah menunjukkan dominasi di segmen mobil listrik... Indonesia masih ribut ngurusi Kadrun... maka enyahkan kadrun mulai 2024 ini	-1	0
Volvo akan bawa jajaran model XC dan 2 mobil listrik murni	0	0
Ternyata Segini Harga Baterai Mobil Listrik Pakar Otomotif Nilai Harga Bekas Wuling Air EV Bakal Terjun Bebas	0	0
@CNNIndonesia Mandatory Spending buat Kesehatan di kurangi eh malah subsidi Mobil Listrik yang bikin macet	0	0
@jundialwan_ @ichFaris_ perlu hitung jg effort utk menahan hedonic treadmill misal jd kepikiran punya heater nambah ac kulkas 7 pintu motor mobil listrik	0	0
Perbandingannya cukup jauh dengan asumsi tarif listrik sebesar Rp1.699 per Kwh hanya diperlukan sekitar Rp2.500 untuk sepeda motor listrik menempuh jarak 50 km dan 10 km untuk mobil listrik	1	0

Analisis Hasil Evaluasi Model

Evaluasi model dilakukan untuk menilai sejauh mana model mampu mengklasifikasikan sentimen yang terkandung dalam data teks secara akurat. Metrik evaluasi yang digunakan mencakup *precision*, *recall*, dan *F1-score* untuk masing-masing kelas sentimen: positif, negatif, dan netral. Model dilatih

menggunakan representasi fitur berbasis Word2Vec dan diuji dalam dua kondisi, yaitu sebelum dan sesudah dilakukan oversampling dengan SMOTE. Pada tahap awal, model dilatih menggunakan data asli yang memiliki distribusi kelas tidak seimbang, di mana jumlah tweet dengan label "netral" jauh lebih dominan dibanding kelas "positif" dan "negatif". Hasil evaluasi menunjukkan bahwa model hanya memprediksi data uji ke dalam kelas netral, sedangkan kelas positif dan negatif sama sekali tidak teridentifikasi. Hal ini menyebabkan nilai precision, recall, dan F1-score untuk kedua kelas tersebut adalah 0.00, yang menunjukkan bahwa model mengalami bias terhadap kelas mayoritas. Sebagai upaya perbaikan, dilakukan teknik oversampling menggunakan metode SMOTE untuk menyeimbangkan jumlah data antar kelas. Teknik ini secara teoritis memungkinkan model memperoleh representasi yang lebih setara dari setiap kelas sentimen. Namun, setelah dilakukan pelatihan ulang terhadap model dengan data yang telah di oversampling, hasil evaluasi tetap menunjukkan bahwa kelas positif dan negatif tidak berhasil diprediksi. Model masih menghasilkan prediksi satu kelas (netral) untuk seluruh data uji, sehingga nilai metrik untuk kelas lainnya tetap 0.00. Kondisi ini menunjukkan bahwa penyebab utama kegagalan klasifikasi bukan hanya berasal dari ketidakseimbangan data, tetapi juga dari beberapa faktor lain, seperti kualitas representasi fitur teks vektor rata-rata Word2Vec mungkin kurang menangkap konteks dan polaritas sentimen. Dengan demikian, untuk meningkatkan kinerja klasifikasi sentimen, diperlukan pendekatan tambahan seperti teknik representasi fitur yang lebih kontekstual seperti TF-IDF, agar informasi penting dalam teks tetap terjaga.

5. Kesimpulan

Tujuan dari penelitian ini adalah untuk melakukan analisis sentimen masyarakat terhadap mobil listrik di Indonesia dengan memanfaatkan data dari platform media sosial X (sebelumnya dikenal sebagai Twitter), menggunakan metode Long Short-Term Memory (LSTM). Dataset yang digunakan terdiri dari 10.283 tweet yang telah divalidasi oleh lima responden, kemudian diklasifikasikan pada tiga kategori sentimen, yaitu positif, netral, dan negatif.

Berdasarkan hasil pengujian yang di proses menggunakan tiga skenario pembagian data (70:30, 80:20, dan 90:10), baik sebelum maupun sesudah dilakukan proses *oversampling* menggunakan metode SMOTE, diketahui bahwa model LSTM memiliki kemampuan klasifikasi yang terbatas. Model menunjukkan kinerja yang baik hanya pada kelas sentimen netral dengan *f1-score* mencapai 71%–72%, namun gagal mengklasifikasikan kelas sentimen positif dan negatif secara akurat, yang ditunjukkan oleh nilai *precision*, *recall*, dan *f1-score* sebesar 0%. Nilai akurasi secara keseluruhan juga cenderung stagnan di kisaran 55%–56%.

Secara keseluruhan, meskipun model LSTM dapat mempelajari pola sentimen netral dengan baik, performanya belum optimal untuk klasifikasi sentimen multi-kelas secara menyeluruh. Keterbatasan performa model ini disebabkan oleh ketidakseimbangan data antar kelas sebelum dilakukan *oversampling*, serta kurangnya efektivitas model LSTM dalam menangkap pola sentimen dari data teks informal seperti *tweet*, meskipun distribusi kelas telah diseimbangkan. Oleh karena itu, dapat disimpulkan bahwa metode LSTM dalam kondisi dan

parameter yang digunakan pada penelitian ini belum mampu memberikan hasil klasifikasi yang optimal untuk sentimen positif dan negatif.

Referensi

- [1] A. Karimah and G. Dwilestari, "ANALISIS SENTIMEN KOMENTAR VIDEO MOBIL LISTRIK DI PLATFORM YOUTUBE DENGAN METODE NAIVE BAYES," 2024.
- [2] W. Ningsih, B. Alfianda, R. Rahmaddeni, and D. Wulandari, "Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada Penggunaan Mobil Listrik di Indonesia," 2024.
- [3] N. Dienwati, G. Dwilestari, and E. Tohidi, "PENERAPAN ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR UNTUK ANALISIS SENTIMEN YOUTUBE MENGENAI INTENSIF MOBIL LISTRIK," 2023.
- [4] T. Prasetyo, A. Adhyaksa Waskita, T. Taryo, J. Raya Puspitak, and K. Pamulang Kota Tangerang Selatan, "Teguh Prasetyo, Arya Adhyaksa Waskita, Taswanda Taryo Analisis Sentimen Pengguna Seputar Kendaraan Listrik Di Twitter Dengan Penerapan Algoritma Naïve Bayes, KNN, dan Decision Tree untuk Klasifikasi." 2025.
- [5] S. J. Pipin, H. Kurniawan, | Jurnal, and S. Mikroskil, "Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM," 2022.
- [6] S. Alfataah and N. Cahyono, "SENTIMEN ANALISIS PENGGUNA TWITTER TERHADAP PROVIDER XL AXIATA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE," 2025.
- [7] E.T. Arifin, Jondri "Prediksi Retweet Menggunakan Fitur User based dan Content Based dengan Metode Klasifikasi ANN-GA."2022.
- [8] M. Z. Rahman, Y. A. Sari, and N. Yudidistra, "Analisis Sentimen Tweet COVID-19 menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM)," 2021.
- [9] M. N. Romzi, I, Atastina "Analisis Sentiment Judul Berita Ekonomi Terhadap Index Harga Saham Gabungan Menggunakan Metode Long short-Term Memory." 2023.
- [10] F. A. Fadhillah, "ANALISIS SENTIMEN MASYARAKAT TERHADAP PENGGUNAAN KENDARAAN LISTRIK PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES."2024.
- [11] M. Nouval, D. Ramadhan, and A. Gunaryati, "KLASIFIKASI SENTIMEN PUBLIK TERHADAP KEBIJAKAN KENDARAAN LISTRIK MENGGUNAKAN ALGORITMA NAÏVE BAYES DAN SUPPORT VECTOR MACHINE," 2025.
- [12] M. Farid, N. S. Ferdiana Kusuma, J. Ngagel, and J. Selatan, "Analisis Sentimen pada Media Sosial Twitter Terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Berbasis Deep Learning," 2022.
- Muthmainnah, Naufal Azmi Verdikha, Fendy Yulianto "KLASIFIKASI ULASAN APLIKASI SIREKAP 2024 DENGAN EKSTRAKSI FITUR WORD2VEC DAN METODE SUPPORT VECTOR MACHINE (SVM)", 2025.