

Analisis Sentimen Kendaraan Listrik Pada Twitter Menggunakan Metode Long Short Term Memory

Dian Agus Prawinata

Universitas Pembangunan Nasional Veteran Jawa Timur

Email: 19081010074@student.upnjatim.ac.id

Ani Dijah Rahajoe

Universitas Pembangunan Nasional Veteran Jawa Timur

Email: anidijah.if@upnjatim.ac.id

I Gede Susrama Mas Diyasa

Universitas Pembangunan Nasional Veteran Jawa Timur

Email: igsusrama.if@upnjatim.ac.id

Alamat: Jl. Rungkut Madya No1, Kec. Gn. Anyar, Surabaya, Jawa Timur, Indonesia

Abstract. *In facing the increasing awareness of environmental impact, electric vehicles have become a primary focus in the global automotive industry. With the advancement of technology and the growing need for eco-friendly solutions, the evaluation of public sentiment towards electric vehicles becomes highly relevant. This research aims to analyze opinions expressed on Twitter regarding the use of electric vehicles using the Long Short Term Memory (LSTM) classification method. Utilizing a dataset of 30,000 entries, this study applies the LSTM algorithm to classify sentiment in tweets. Four different scenarios are tested, involving combinations of Continuous Bag of Words (CBOW) and Skip-Gram feature extraction methods, as well as data split percentages of 80:20 and 70:30. The research results demonstrate high accuracy levels across all scenarios, ranging from 85.16% to 85.9%. These findings indicate the effectiveness of sentiment analysis in gauging public perspectives on the use of electric vehicles. This study makes a significant contribution to understanding public sentiment related to electric vehicles based on Twitter data while highlighting the application of sentiment analysis techniques in the context of electric vehicle usage.*

Keywords: *Sentiment Analysis, Electric Vehicles, Twitter, LSTM.*

Abstrak. Dalam menghadapi peningkatan kesadaran terhadap dampak lingkungan, kendaraan listrik menjadi fokus utama dalam industri otomotif global. Seiring dengan perkembangan teknologi dan meningkatnya kebutuhan akan solusi ramah lingkungan, evaluasi sentimen publik terhadap kendaraan listrik menjadi sangat relevan. Penelitian ini bertujuan untuk menganalisis opini yang diungkapkan di Twitter mengenai penggunaan kendaraan listrik, menggunakan metode klasifikasi *Long Short Term Memory* (LSTM). Menggunakan dataset sebanyak 30.000 data, penelitian ini menerapkan algoritma LSTM untuk mengklasifikasikan sentimen tweet. Empat skenario berbeda diuji, melibatkan kombinasi metode ekstraksi fitur *Continuous Bag of Words* (CBOW) dan *Skip-Gram*, serta persentase pemisahan data 80:20 dan 70:30. Hasil penelitian menunjukkan tingkat akurasi yang tinggi di seluruh skenario, berkisar antara 85,16% hingga 85,9%. Hasil ini menandakan efektivitas analisis sentimen dalam mengukur pandangan masyarakat terhadap penggunaan kendaraan listrik. Studi ini memberikan kontribusi penting dalam memahami sentimen publik terkait kendaraan listrik berdasarkan data Twitter, sambil menyoroti penerapan teknik analisis sentimen dalam konteks penggunaan kendaraan listrik.

Kata kunci: Analisis Sentimen, Kendaraan Listrik, Twitter, LSTM.

LATAR BELAKANG

Kemunculan kendaraan listrik telah menjadi fokus utama dalam sektor transportasi, memanfaatkan listrik sebagai sumber energi utama untuk propulsi (Ahmad et al., 2018). Dibandingkan dengan kendaraan konvensional yang menggunakan bahan bakar minyak, kendaraan listrik menunjukkan tingkat kebisingan mesin yang rendah, tidak menghasilkan

emisi gas buang, dan memiliki efisiensi energi yang tinggi (Song et al., 2016). Pengembangan kendaraan listrik tidak hanya berkontribusi pada pengurangan emisi karbon dari sektor transportasi, tetapi juga memajukan infrastruktur listrik (Ahmad et al., 2018).

Menanggapi pergeseran transformatif ini, pemerintah Indonesia mengeluarkan Peraturan Presiden (PERPRES) Nomor 55 tahun 2019 tentang Kendaraan Listrik, yang mempercepat Program Kendaraan Bermotor Berbasis Baterai untuk jalan. Akibatnya, wacana publik dan reaksi muncul, terutama di Twitter, mencerminkan beragam pendapat mengenai kemunculan kendaraan listrik (Darwis et al., 2021). Penggunaan kendaraan bertenaga listrik yang akan datang pada KTT G20 di Bali lebih memperkuat diskusi melalui hashtag #kendaraanlistrik, meningkatkan keterlibatan publik terhadap mode transportasi ini.

Lanskap digital Indonesia telah menyaksikan kehadiran media sosial yang signifikan, dengan sekitar 167,0 juta pengguna yang aktif terlibat di berbagai platform pada Januari 2023, membentuk 60,4% dari total populasi (DIGITAL 2023: INDONESIA, 2023). Di antara platform-platform tersebut, Twitter menonjol, dengan 24,00 juta pengguna di awal 2023, mewakili 8,7% dari populasi negara ini (Kemp, 2023).

Memahami sentimen, persepsi, dan pandangan masyarakat mengenai kendaraan listrik memerlukan analisis percakapan di Twitter. Analisis sentimen, suatu area studi yang berfokus pada mengevaluasi opini, emosi, dan sikap dari teks atau ucapan, berfungsi sebagai alat yang kuat dalam hal ini (Bing Liu, 2012).

Dalam konteks ini, studi seperti "Penerapan Analisis Sentimen dan Naïve Bayes terhadap Opini tentang Penggunaan Kendaraan Listrik di Twitter" oleh Aditta Agustian, Tukino, dan Fitria Nurapriani, menggunakan metode Naïve Bayes, mencapai tingkat akurasi 80%. Sebaliknya, "Analisis Sentimen atas Ulasan Novel dengan Menggunakan Metode Long Short-Term Memory" oleh Muh Amin Nurrohmat dan Azhari SN, dengan menerapkan *Deep Learning* - khususnya, metode *Long Short-Term Memory* (LSTM) - mengungguli Naïve Bayes dalam studi kasus serupa.

LSTM, suatu varian dari *Recurrent Neural Networks* (RNNs), unggul dalam mempertahankan informasi jangka panjang dengan menggabungkan sel memori dan tiga pintu - input, lupakan, dan output - mengatasi tantangan yang dihadapi oleh RNN tradisional (S. Hochreiter dan J. Uergen Schmidhuber, 1997).

Beberapa studi telah menggunakan berbagai algoritma seperti Support Vector Machine (SVM) dan ekstraksi fitur TF-IDF, mencapai kinerja klasifikasi 87% (Muhammad Ilham Alhari et al., 2022). Selain itu, sebuah studi oleh Daniel Haryadi dan Gede Putra Kusuma, dengan

menggunakan LSTM dan dataset yang signifikan, mencatat tingkat akurasi 99,167% beserta nilai presisi, recall, dan F1 yang mengesankan (2019).

Dengan latar belakang ini, penelitian ini memanfaatkan Twitter sebagai sumber data dan mengimplementasikan *Deep Learning* - khususnya, metode LSTM - untuk menganalisis sentimen mengenai "kendaraan listrik" di platform tersebut.

KAJIAN TEORITIS

Penelitian Terdahulu

Penelitian oleh Daniel Haryadi dan Gede Putra Kusuma pada tahun 2019 menggunakan dataset Twitter dengan 980,549 data training dan 144,160 data testing, diambil melalui API Twitter. Proses preprocessing melibatkan penghapusan tanda baca dan perubahan kata menjadi huruf kecil. Dalam klasifikasi menggunakan LSTM dan *Nested LSTM*, Model *Nested LSTM* mencapai akurasi tertinggi 99,167%, mendekati akurasi LSTM sebesar 99,154%. Walaupun begitu, model LSTM secara konsisten memberikan performa terbaik dengan presisi rata-rata 99,22%, recall rata-rata 98,86%, dan skor f1 99,04%.

Pada Penelitian yang dilakukan oleh Naila Aslam, Furqan Rustam, Ernesto Lee, Patrick Bernard Washington dan Imran Ashraf pada tahun 2022. Hasil yang diperoleh dari klasifikasi menggunakan LSTM-GRU pada dataset twitter adalah Model mencapai kinerja tertinggi untuk analisis sentimen dengan skor akurasi 0,99 dan presisi dan recall masing-masing 0,99 dan 0,98.

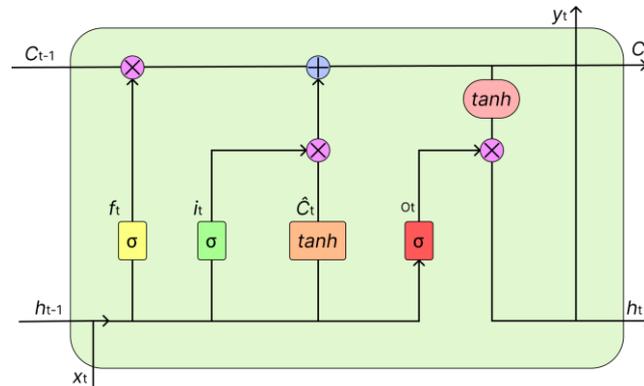
Analisis Sentimen

Analisis sentimen adalah sebuah bidang ilmu yang bertujuan untuk menganalisis opini, evaluasi, sentimen, sikap, dan emosi yang terungkap dalam bahasa tulisan seseorang. Dalam praktiknya, analisis sentimen meliputi beberapa aspek berbeda, seperti ekstraksi opini, penambangan sentimen, analisis subjektivitas, analisis dampak, serta analisis emosi. (Bing Liu, 2012).

Long Short Term Memory

Long Short-Term Memory (LSTM) merupakan salah satu variasi dari Recurrent Neural Network (RNN) yang mampu mempertahankan informasi dalam jangka panjang. Dalam LSTM, simpul RNN pada hidden layer digantikan oleh sel LSTM yang dirancang untuk menyimpan informasi masa lalu. LSTM menggunakan tiga gerbang yaitu *input gate*, *forget gate*, dan *output gate* untuk mengontrol penggunaan dan pembaruan informasi masa lalu. Sel memori dan ketiga gerbang tersebut dirancang untuk membaca, menyimpan, dan memperbarui

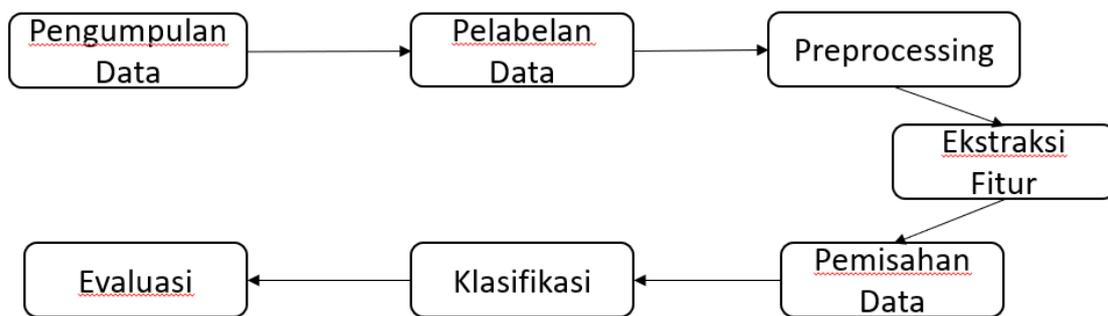
informasi, sehingga LSTM dapat mempelajari pola pada teks yang lebih kompleks (Horchreiter dan Schmidhuber, U., 1997).



Gambar 1. Struktur LSTM

METODE PENELITIAN

Rancangan Penelitian



Gambar 2. Rancangan Penelitian

Penelitian ini melibatkan beberapa tahapan yang mendalam untuk mendapatkan hasil yang akurat dan relevan. Tahapan pertama adalah Pengambilan Data, di mana 30.000 tweet dikumpulkan menggunakan Teknik *Scraping*. Setelah data terkumpul, langkah selanjutnya adalah Pelabelan Data, di mana tweet-tweet tersebut diberi label sebagai Positif, Netral, atau Negatif menggunakan pre-trained IndoBERTBase Model.

Proses selanjutnya adalah *Preprocessing*, di mana data tweet menjalani serangkaian transformasi seperti *Case Folding*, *Filtering*, *Stop Word Removal*, *Tokenisasi*, dan *Stemming*. Tujuan dari tahap ini adalah untuk menghasilkan bentuk data yang lebih sesuai dan mudah diproses dalam langkah-langkah berikutnya.

Ekstraksi Fitur merupakan langkah berikutnya, yang melibatkan penggunaan Metode Word2Vec CBOW dan *Skip-Gram* untuk mengekstrak fitur-fitur penting dari data. Setelah itu, dilakukan Pemisahan Data menggunakan metode random state untuk memastikan keberagaman data yang digunakan dalam pengembangan dan pengujian model.

Langkah selanjutnya adalah Klasifikasi, di mana metode LSTM digunakan untuk mengkategorikan tweet berdasarkan label yang telah ditentukan sebelumnya. Terakhir, dilakukan Evaluasi kinerja model, yang mencakup pengukuran Akurasi, Presisi, *F1-score*, dan recall.

Dataset

Metode penelitian ini memanfaatkan dataset yang terdiri dari data tweet berbahasa Indonesia yang mencakup kata kunci "kendaraan listrik". Sebanyak 30.000 data tweet telah dihimpun untuk analisis, yang melibatkan pengambilan data dari tanggal 1 November 2022 hingga 10 Maret 2023. Dalam pengumpulan data, tidak dilakukan penyaringan terhadap tweet yang mungkin diposting oleh akun media sosial yang memiliki peran dalam menyebarkan informasi atau melakukan promosi, sehingga memastikan keberagaman dan representativitas data yang digunakan.

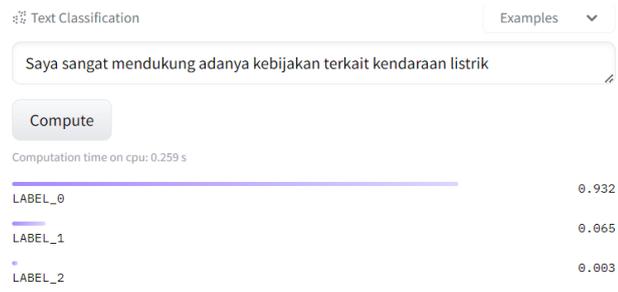
Dataset ini kemudian mengalami serangkaian tahapan *Preprocessing* yang melibatkan *Case Folding*, *Filtering*, *Stop Word Removal*, *Tokenisasi*, dan *Stemming*. Tujuannya adalah untuk memastikan data yang bersih dan siap diproses dalam analisis selanjutnya.

Pelabelan Dataset

Selanjutnya, data tweet dilakukan pelabelan data menjadi 3 kelas, yaitu positif, negatif, dan netral. Pelabelan dilakukan dengan memanfaatkan model *pre-trained* IndoBERTBASE yang menggunakan Prosa Sentiment Dataset yang berisi komentar dan *review* dari berbagai media sosial.

IndoBERT adalah model bahasa canggih yang telah dilatih khusus untuk memproses bahasa Indonesia dan memiliki kemampuan untuk memahami sentimen dalam teks. Dengan menggunakan teknik-teknik seperti *masked language modeling* (MLM) dan *next sentence prediction* (NSP) selama pelatihan, IndoBERT dapat mengenali sentimen seperti positif, negatif, dan netral dalam kalimat atau dokumen bahasa Indonesia. Dengan kata lain, model ini dapat digunakan untuk melabeli teks berdasarkan sentimen yang terkandung dalamnya.

Model IndoBERTBASE menggunakan batch size 256 dan learning rate $2e-5$. Model dilatih menggunakan *Masked Language Modeling* (MLM). Model IndoBERTBASE dilatih pada TPUv3-8 dengan maksimal sequence length 128. Pelatihan ini memakan waktu sekitar 35 jam dan menghasilkan performa skor f1 sebesar 87,73%. Penjelasan lebih lengkap mengenai IndoBERT dapat dilihat pada jurnal "*IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding*". (Wilie et al, 2020).



Gambar 3. Contoh pelabelan menggunakan Indobert

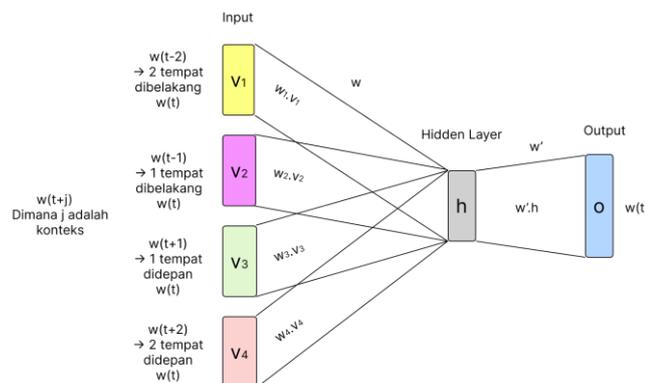
Dalam Gambar 3, terlihat hasil pelabelan menggunakan IndoBERT. Label_0, yang menunjukkan sentimen positif, memiliki akurasi sebesar 0.932. Sementara itu, Label_1, yang menunjukkan sentimen netral, memiliki akurasi sebesar 0.065, dan Label_2, yang menunjukkan sentimen negatif, memiliki akurasi sebesar 0.003. Oleh karena itu, berdasarkan hasil pelabelan tersebut, dapat disimpulkan bahwa sentimen yang terkandung adalah positif.

Ekstraksi Fitur Word2Vec

Proses pembentukan vektor kata menggunakan word2vec melibatkan teknik penghasilan representasi vektor kata yang dikembangkan oleh Google. Dengan menggunakan algoritma word2vec, setiap kata dalam suatu teks atau korpus dapat direpresentasikan sebagai vektor numerik dalam ruang N dimensi, di mana N merupakan ukuran dimensi yang telah ditentukan sebelumnya. (Mikolov et al, 2013)

1. Continuous Bag of Word (CBOW)

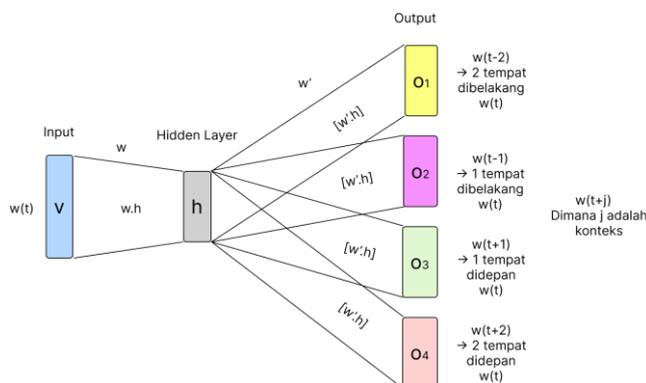
CBOW berusaha untuk memprediksi kata target berdasarkan pada konteks kata-kata di sekitarnya. Ide dasarnya adalah dengan menggunakan beberapa kata konteks sekitar kata target, CBOW berusaha untuk memprediksi kata tersebut. Model CBOW bekerja dengan mengambil kata-kata konteks sebagai input, kemudian mencoba memprediksi kata target. Dengan cara ini, CBOW dapat belajar vektor representasi kata yang mencerminkan hubungan antara kata-kata berdasarkan konteks yang sering muncul bersama.



Gambar 4. Ilustrasi CBOW

2. Skip-Gram

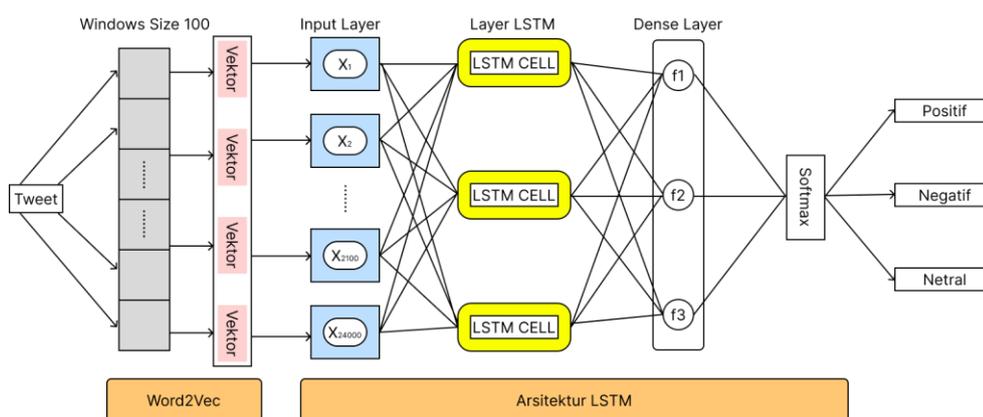
Skip-Gram bekerja sebaliknya dengan CBOW. Model *Skip-Gram* mencoba memprediksi kata-kata konteks (beberapa kata di sekitar) berdasarkan kata target. Dengan kata lain, *Skip-Gram* menggunakan kata target sebagai input dan mencoba memprediksi kata-kata yang mungkin muncul di sekitarnya. *Skip-Gram* secara efektif belajar untuk membuat vektor representasi kata yang lebih baik dengan mengenali kata-kata yang sering muncul bersama-sama.



Gambar 5. Ilustrasi Skip-Gram

LSTM

Dalam konteks klasifikasi data, LSTM dapat digunakan untuk memproses dan menganalisis teks atau urutan data lainnya, serta melakukan klasifikasi berdasarkan pola-pola yang ada dalam urutan data tersebut. Misalnya, dalam analisis sentimen teks pada penelitian ini, LSTM dapat mempelajari pola-pola teks yang berkaitan dengan sentimen positif, negatif, atau netral.



Gambar 6. Tahapan Klasifikasi Data

Model LSTM yang digunakan adalah *Vanilla LSTM*. Ini adalah model dasar dari LSTM yang terdiri dari gerbang *input*, gerbang lupa, dan gerbang *output*. Model ini memiliki sel memori yang memungkinkan informasi untuk mengalir ke selanjutnya atau dilupakan.

Confusion Matrix

Pada tahapan pengujian sistem ini, dilakukan analisis dan evaluasi kinerja sistem yang telah dikembangkan menggunakan metode Confusion Matrix. Confusion Matrix adalah sebuah matriks yang digunakan untuk menggambarkan performa dari suatu sistem klasifikasi berbasis data. Terdapat empat kategori penting yang terdiri dari *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) yang membentuk elemen-elemen dalam Confusion Matrix.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Semua data yang digunakan dalam penelitian ini diperoleh dari Twitter dengan menggunakan *library Snsrape*. *Library* ini digunakan untuk mengambil tweet beserta waktunya. Data yang diambil merupakan data tanggal tweet dan isi tweet berbahasa Indonesia mulai dari tanggal 1 November 2022 hingga 10 Maret 2023 dengan jumlah data sebesar 30.000 data. Contoh hasil pengumpulan data dapat dilihat pada tabel 1.

Tabel 1. Pengumpulan Data

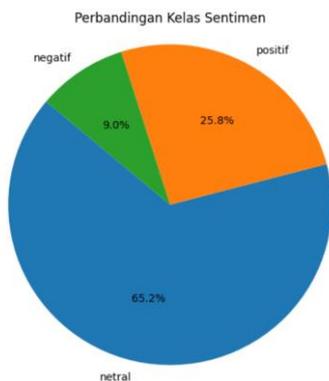
Tweets
@detikcom Gubernur siapapun ga bakal bs bikin semua ruas jalanan Jakarta lega. Ya memang jumlah kendaraan pribadi udah terlalu banyak. Literally TERLALU BANYAK. Udah gitu malah dibanyakin LCGC. 🙄 Makin parahnya lagi, peralihan mobil listrik disubsidi 🙄 Ya gimana ga makin mbludak tu mobil Electrizen, PLN terus memberikan kemudahan bagi pengguna kendaraan listrik atau <i>Electric Vehicle</i> (EV), salah satunya melalui layanan pengisian daya kendaraan listrik di rumah atau <i>home charging</i> . #HomeChargingPLN #ElectricVehicle #PLNMobile https://t.co/MSGr3X64cM
#MostPopuler Subsidi Bikin Harga Kendaraan Listrik Jadi Lebih Terjangkau https://t.co/w1rwqwK0iK

Pelabelan Data

Tabel 2. Jumlah Data Setiap Kelas

Label Sentimen	Jumlah Data
Netral	19557
Positif	7745
Negatif	2698

Jumlah data yang terklasifikasi sebagai 'netral' berjumlah 19.557 data, sedangkan data dengan sentimen 'positif' memiliki jumlah 7.745 data. Sementara itu, data yang masuk ke dalam kategori 'negatif' memiliki jumlah yang lebih sedikit, yaitu sebanyak 2,698 data. Perbandingan presentase tiap kelas dapat dilihat pada gambar 7.



Gambar 7. Diagram Lingkaran Label Sentimen

Preprocessing Data

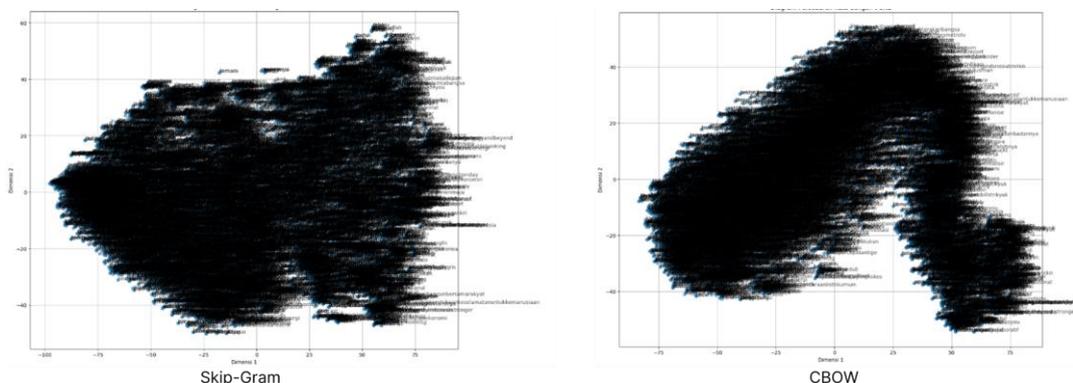
Proses ini dilakukan untuk mempersiapkan data agar siap masuk ke tahap pelatihan. Data tersebut melalui beberapa proses antara lain *case folding*, *cleaning*, normaliasi kata alay, *stop word removal*, *stemming* dan *tokenizing*. Hasil Data telah dilakukan *preprocessing* dapat dilihat pada tabel 3.

Tabel 3. Hasil Preprocessing

Tweets
gubernur,bikin,ruas,jalan,jakarta,lega,ya,kendara,pribadi,literally,dibanyakin,lcgc,parah,alih,mobil,listrik,subsidi,ya,mbludak,mobil
electrizen,usaha,listrik,negara,mudah,guna,kendara,listrik,ev,salah,satu,layan,isi,daya,kendara,listrik,rumah,homechargingpln,electricvehicle,plnmobile
mostpopuler,subsidi,bikin,harga,kendara,listrik,jangkau

Ekstraksi Fitur

Teknik Word2Vec memanfaatkan jaringan saraf tiruan (neural networks) untuk mempelajari representasi distribusional kata-kata dalam suatu korpus teks. Berdasarkan data yang dimasukkan terdapat 14.356 kata unik. Pada Penelitian ini digunakan 2 skenario metode Word2Vec yaitu, *Continuous Bag of Words* (CBOW) dan *Skip-Gram*. Pada Gambar 8 dapat dilihat persebaran kata yang direduksi menjadi vektor 2 dimensi.



Gambar 8. Diagram Persebaran Kata

Pemisahan Data

Pemisahan data menjadi dataset pelatihan (*training dataset*) dan dataset pengujian (*test dataset*) dilakukan untuk menguji kinerja model yang dikembangkan pada data yang tidak pernah dilihat sebelumnya oleh model tersebut. Pada penelitian ini akan dilakukan pemisahan data dalam dua skenario.

1. 80% data *training* dan 20 % data *testing*

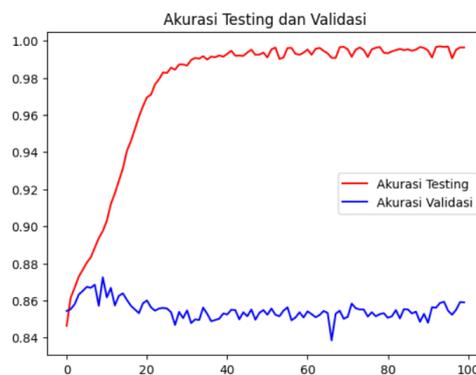
Dari total 30.000 data, data tersebut dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Data *training* mencakup 15.636 data netral, 6.232 data positif dan 2.132 data negatif. Sementara itu, data *testing* terdiri dari 3.921 data netral, 1.513 data positif dan 566 data negatif.

2. 70% data *training* dan 30% data *testing*

Dari total 30.000 data, data tersebut dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Data *training* mencakup 13.647 data netral, 5.473 data positif, dan 1.880 data negatif. Sementara itu, data *testing* terdiri dari 5.910 data netral, 2.272 data positif, dan 818 data negatif.

Hasil Klasifikasi LSTM

1. *Skip-Gram* Rasio Pemisahan Data 80:20



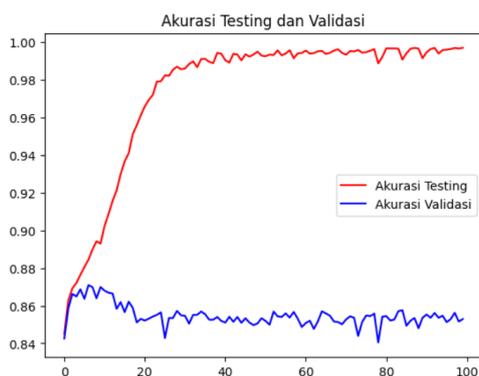
Gambar 9. Pelatihan LSTM dengan *Skip-Gram* Rasio Pemisahan Data 80:20

Gambar 9 merupakan hasil pelatihan algoritma LSTM dengan membatasi pelatihan hanya sampai 100 *epoch*. Hasil perhitungan pengujian sistem menggunakan confusion matrix dapat dilihat pada tabel 4.

Tabel 4. Confusion matrix skenario *Skip-Gram* Rasio Pemisahan Data 80:20

Aktual	Prediksi		
	3606	135	180
	184	334	48
	239	60	1214

2. Skip-Gram Rasio Pemisahan Data 70:30



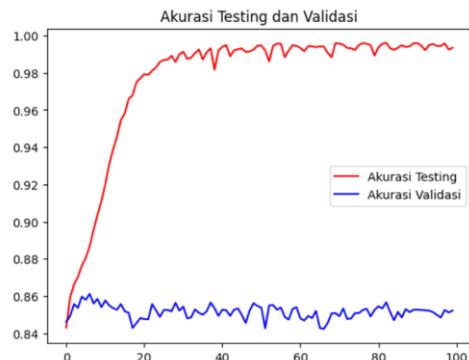
Gambar 10. Pelatihan LSTM dengan Skip-Gram Rasio Pemisahan Data 70:30

Gambar 10 merupakan hasil pelatihan algoritma LSTM dengan membatasi pelatihan hanya sampai 100 epoch. Hasil perhitungan pengujian sistem menggunakan confusion matrix dapat dilihat pada tabel 5.

Tabel 5. Confussion matrix skenario Skip-Gram Rasio Pemisahan Data 70:30

Aktual	Prediksi		
	5356	227	327
257	479	82	
346	85	1841	

3. CBOW Rasio Pemisahan Data 80:20



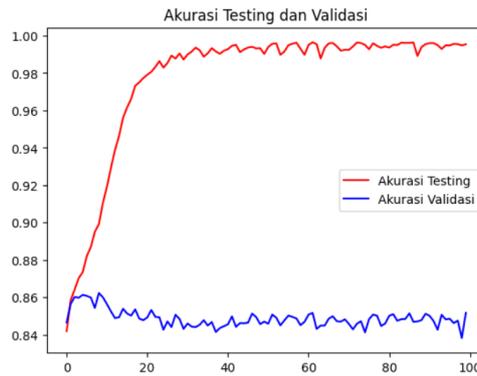
Gambar 11. Pelatihan LSTM dengan CBOW Rasio Pemisahan Data 80:20

Gambar 11 merupakan hasil pelatihan algoritma LSTM dengan membatasi pelatihan hanya sampai 100 epoch. Hasil perhitungan pengujian sistem menggunakan confusion matrix dapat dilihat pada tabel 6.

Tabel 6. Confussion matrix skenario Skip-Gram Rasio Pemisahan Data 80:20

Aktual	Prediksi		
	3576	142	203
192	300	74	
209	67	1237	

4. CBOW Rasio Pemisahan Data 70:30



Gambar 12. Pelatihan LSTM dengan CBOW Rasio Pemisahan Data 70:30

Gambar 12 merupakan hasil pelatihan algoritma LSTM dengan membatasi pelatihan hanya sampai 100 epoch. Hasil perhitungan pengujian sistem menggunakan confusion matrix dapat dilihat pada tabel 7.

Tabel 7. Confusion matrix skenario Skip-Gram Rasio Pemisahan Data 70:30

Aktual	Prediksi		
	5422	227	261
296	440	82	
381	88	1803	

Hasil Evaluasi menggunakan confusion matrix

Tabel 8. Perbandingan hasil evaluasi menggunakan confusion matrix

Skenario	Akurasi	Presisi	Recall	F1-Score
Skenario 1	0.859	0.8567	0.859	0.8578
Skenario 2	0.8528	0.8518	0.8528	0.8523
Skenario 3	0.8521	0.8492	0.8521	0.8506
Skenario 4	0.8516	0.8488	0.8516	0.8502

Skenario pengujian pertama dengan menggunakan Ekstraksi Fitur *Skip-Gram* dan Pemisahan data berupa 80% data *training* dan 20 % data *testing* menghasilkan akurasi tertinggi sebesar 0.859 atau 85.9 %, Presisi sebesar 0.8567 atau 85.67%, Recall sebesar 0.859 atau 85.9 % dan F1 Score sebesar 0.8578 atau 85.78%.

KESIMPULAN DAN SARAN

Kesimpulan

Kesimpulan dari penelitian ini menunjukkan bahwa Metode *Long Short-Term Memory* (LSTM) efektif digunakan untuk melakukan analisis sentimen terkait kendaraan listrik di platform Twitter. Pengujian dengan berbagai skenario menunjukkan bahwa skenario pertama, yang melibatkan Ekstraksi Fitur *Skip-Gram* dan Pemisahan data dengan rasio 80% data

training dan 20% data *testing*, menghasilkan kinerja tertinggi dengan akurasi sebesar 82.8%, presisi 82.68%, recall 82.8%, dan F1 *Score* 82.74%.

Saran

Berdasarkan hasil dan kesimpulan dari penelitian ini, saran untuk penelitian selanjutnya dapat difokuskan pada peningkatan kualitas analisis sentimen dengan beberapa langkah, seperti menyeimbangkan jumlah data tiap kelas menggunakan metode seperti Synthetic Minority Over-sampling Technique (SMOTE). Selain itu, eksplorasi penggunaan model LSTM yang lebih kompleks dapat menjadi langkah berikutnya untuk meningkatkan daya prediktif dan keakuratan analisis sentimen terhadap tweet berbahasa Indonesia mengenai kendaraan listrik di media sosial. Keseluruhan, upaya-upaya ini diharapkan dapat meningkatkan relevansi dan kehandalan analisis sentimen dalam konteks kendaraan listrik di era digital ini.

DAFTAR REFERENSI

- Agustian, T., Tukiro, T., & Nurapriani, F. (2022). Analisis Sentimen, Text Mining Penerapan Analisis Sentimen Dan Naive Bayes Terhadap Opini Penggunaan Kendaraan Listrik Di Twitter. *TIKA*, 7(3), 243-249. doi: <https://doi.org/10.51179/tika.v7i3.1550>.
- Ahmad, A., Khan, Z., Alam, M., & Khateeb, S. (2018). A Review of The Electric Vehicle Charging Techniques, Standards, Progression and Evolution of EV Technologies in German. *Smart Science*, 6(1), 36-53.
- Alhari, M. I., Pratiwi, O. N., & Lubis, M. (2022). Sentiment Analysis of The Public Perspective on Electric Cars in Indonesia Using Support Vector Machine Algorithm. In *Proceedings of the 2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*. DOI: 10.1109/ICSINTESA56431.2022.10041604.
- Aslam, N., Rustam, F., Lee, E., Ashraf, A., & Bernard, P. (2022). Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model. *IEEE Access*. doi: 10.1109/ACCESS.2022.3165621.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter BMKG Nasional. *Jurnal Tekno Kompak*, 15(1), 131. doi: 10.33365/jtk.v15i1.744.
- Deng, L., & Yu, D. (2012). *Deep Learning: Methods and Applications*. *Foundation and Trends in Signal Process.*, 7(3–4), 197–387, 2014. doi: 10.1561/20000000039.
- Haryadi, D., & Kusuma, G. P. (2019). Emotion Detection in Text using Nested Long Short-Term Memory. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 10(6).
- Hochreiter, S., & Schmidhuber, U. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.
- JDIH BPK RI. (n.d.). Percepatan Program Kendaraan Bermotor Listrik Berbasis Baterai (Battery Electric Vehicle) untuk Transportasi Jalan. URL: <https://peraturan.bpk.go.id/Home/Details/116973/perpres-no-55-tahun-2019>.

- Kemp, S. (2023). Digital 2023: Indonesia. <URL: <https://datareportal.com/reports/digital-2023-indonesia>>.
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers.
- Liu, B., Hu, M., & Cheng, J. (2005). Opinion Observer: Analyzing and Comparing Opinions on the Web. In Proceedings of the 14th International World Wide Web Conference (WWW-2005), May 10-14, Chiba, Japan.
- Nurrohmat, M. A., & Sn, A. (Jul. 2019). Sentiment Analysis of Novel Review Using *Long Short-Term Memory* Method. IJCCS, 13(3), 209. doi: 10.22146/ijccs.41236.
- Olah, C. (2015). Understanding LSTM Networks. URL: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- Song, Y., Li, J., Ji, G., & Xue, Z. (2016). Study on The Typical Mode EV Charging and Battery Swap Infrastructure Interconnecting to Power Grid. China International Conference on Electric Distribution.
- Vrigazova, B. (2021). The Proportion for Splitting Data into *Training* and Test Set for the Bootstrap in Classification Problems. Business Systems Research Journal. DOI: <https://doi.org/10.2478/bsrj-2021-0015>.
- Wahid, D. H., & Azhari, S. N. (2016). Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 10(2), 207-218.
- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. In Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing.