

ANALISIS SENTIMEN PADA KURIKULUM MERDEKA MENGGUNAKAN WORD2VEC DAN ALGORITMA LONG SHORT-TERM MEMORY

Octhavia Alin*¹⁾, Ema Utami²⁾

1. Universitas Amikom Yogyakarta, Indonesia
2. Universitas Amikom Yogyakarta, Indonesia

Article Info

Kata Kunci: Kurikulum Merdeka; LSTM; Sentimen Analisis; Word2Vec;

Keywords: Kurikulum Merdeka; LSTM; Sentimen Analisis; Word2Vec;

Article history:

Received 24 August 2024

Revised 6 September 2024

Accepted 3 October 2024

Available online 1 September 2025

DOI :

<https://doi.org/10.29100/jipi.v10i3.6516>

* Corresponding author.

Corresponding Author

E-mail address:

alin.octhavia@students.amikom.ac.id

ABSTRAK

Kurikulum Merdeka, sebagai penyempurnaan dari Kurikulum 2013, telah memicu berbagai opini publik sejak diterapkan pada Tahun Ajaran 2022/2023. Opini-opini ini, yang dibagikan melalui platform media sosial seperti Twitter dan YouTube, mencerminkan pandangan beragam dari masyarakat. Penelitian ini menganalisis sentimen terhadap Kurikulum Merdeka menggunakan algoritma LSTM yang dipadukan dengan Word2Vec. Dataset yang digunakan terdiri dari 23.577 teks dari Twitter dan YouTube, yang dikategorikan menjadi sentimen negatif, positif, dan netral. Penelitian ini membandingkan dua arsitektur *Word2Vec*. *Skip-gram* dan *Continuous Bag of Words (CBOW)*, serta mengeksplorasi pengaruh *optimizer* (Adam dan RMSprop) dan *learning rate* terhadap kinerja model. Hasil penelitian menunjukkan bahwa arsitektur CBOW, yang dikombinasikan dengan *optimizer* Adam dan *learning rate* 0.001, memberikan hasil yang paling akurat dan stabil, dengan akurasi mencapai 97% dan tingkat *loss* yang rendah. *Skip-Gram* lebih sensitif terhadap perubahan parameter, yang menyebabkan ketidakstabilan kinerja. Penelitian ini menegaskan efektivitas kombinasi *Word2Vec* dan LSTM untuk analisis sentimen, dengan CBOW dan *optimizer* Adam sebagai konfigurasi yang paling optimal.

ABSTRACT

The Merdeka Curriculum, an enhancement of the 2013 Curriculum, has sparked various public opinions since its implementation in the 2022/2023 academic year. These opinions, shared through social media platforms such as Twitter and YouTube, reflect diverse views from the public. This study analyzes sentiment towards the Merdeka Curriculum using the LSTM algorithm combined with Word2Vec. The dataset used consists of 23,577 texts from Twitter and YouTube, categorized into negative, positive, and neutral sentiments. This research compares two Word2Vec architectures, Skip-gram and Continuous Bag of Words (CBOW), and explores the influence of optimizers (Adam and RMSprop) and learning rates on model performance. The results show that the CBOW architecture, combined with the Adam optimizer and a learning rate of 0.001, provides the most accurate and stable outcomes, with accuracy reaching up to 97% and low loss rates. Skip-gram is more sensitive to parameter changes, leading to instability in performance. This study confirms the effectiveness of combining Word2Vec and LSTM for sentiment analysis, with CBOW and the Adam optimizer being the most optimal configuration.

I. PENDAHULUAN

KURIKULUM Merdeka[1] merupakan versi yang telah disempurnakan dari Kurikulum 2013, yang dirancang untuk menyesuaikan dengan kondisi masyarakat dan kebutuhan peserta didik setelah pandemi. Kurikulum ini juga dikenal sebagai Kurikulum Paradigma Baru atau Kurikulum 2022, yang memiliki sifat fleksibel serta berfokus pada pengembangan karakter, kompetensi, dan kreativitas. Implementasi Kurikulum Merdeka dimulai pada tahun ajaran 2022/2023 di tingkat pendidikan dasar dan menengah, dengan rencana penerapan di seluruh satuan pendidikan di Indonesia. Saat ini, terdapat tiga kurikulum yang digunakan: Kurikulum 2013, Kurikulum

Merdeka (prototipe), dan Kurikulum Darurat. Peraturan Mendikbud Ristek No. 12 Tahun 2024, Kurikulum Merdeka ditetapkan sebagai kerangka utama untuk semua jenjang pendidikan di Indonesia, dengan implementasi penuh paling lambat tahun ajaran 2026/2027. Kurikulum ini menekankan pada materi-materi inti, memberi ruang lebih bagi guru untuk mengembangkan karakter dan kompetensi siswa, serta mendorong inovasi dalam pengajaran dan layanan pendidikan.

Perubahan Kurikulum Merdeka upaya berkesinambungan untuk meningkatkan kualitas pendidikan. Saat ini, topik ini menjadi bahan diskusi di berbagai lembaga pendidikan. Komentar di Twitter tentang Kurikulum Merdeka mencerminkan kontroversi sejak penerapannya pada Tahun Ajaran 2022/2023, memicu berbagai opini di media sosial[2]. Analisis sentimen terhadap opini ini dapat memberikan wawasan mendalam tentang penerimaan publik dan membantu mengidentifikasi area perbaikan, sehingga keputusan terkait kurikulum dapat diambil dengan lebih baik [3].

Dalam beberapa tahun terakhir, analisis sentimen menjadi metode yang cukup populer, tidak hanya di kalangan akademisi, tetapi juga di sektor bisnis, pemerintah, dan berbagai organisasi lainnya[4]. Juga dikenal sebagai analisis opini, analisis sentimen adalah bagian dari *Natural Language Processing* (NLP) yang berfungsi untuk mengenali dan mengekstrak opini dari teks. Berbagai penelitian telah menerapkan analisis sentimen, seperti yang dilakukan oleh [5] yang meneliti dampak akun-akun palsu di media sosial, khususnya Twitter, terhadap analisis sentimen terkait opini publik mengenai COVID-19 di Indonesia. Artikel tersebut menggunakan pendekatan *machine learning* untuk mengklasifikasikan sentimen positif, negatif, dan netral dari tweet terkait COVID-19, dan menemukan bahwa akun palsu dapat menurunkan performa klasifikasi sentimen, dengan akurasi yang diperoleh sebesar 80,6%.

Dalam mengatasi tantangan dalam analisis sentimen ini, *deep learning* telah diperkenalkan. Sebagai bagian dari *machine learning*, teknologi ini menggunakan algoritma yang mampu menangani masalah kompleks dengan lebih efektif. *Deep learning* bekerja secara berurutan dan memungkinkan komputer untuk mengelola data dalam skala besar dengan sedikit intervensi manusia[6]. LSTM, sebagai bagian dari *deep learning*, dirancang khusus untuk menangani masalah ketergantungan jangka panjang dalam data deret waktu, termasuk dalam analisis sentimen dan proses klasifikasi teks lainnya[7]. Dalam sebuah studi oleh [8], algoritma LSTM digunakan untuk mengembangkan teknik penilaian sentimen terkait ulasan hotel dalam bahasa Indonesia. Studi ini mengintegrasikan model *Word2Vec* dengan LSTM, menggunakan arsitektur *Word2Vec*, dimensi vektor, metode evaluasi, teknik *pooling*, nilai *dropout*, dan *learning rate*. Dari eksperimen yang dilakukan pada 2.500 teks ulasan, studi ini mencapai kinerja terbaik dengan akurasi sebesar 85,96%. *Word2Vec* mengubah kata-kata menjadi vektor numerik untuk memahami konteksnya, membangun vektor berdasarkan distribusi kata [9].

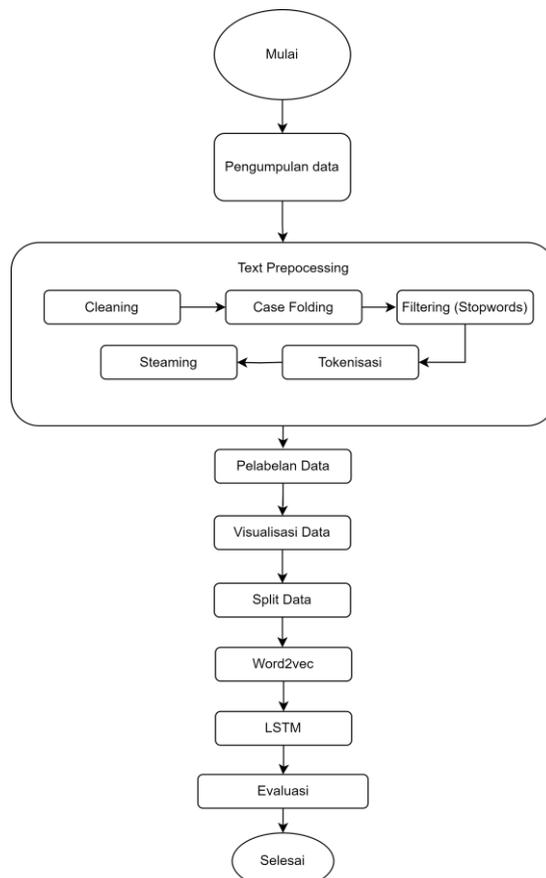
Meskipun *Word2Vec* meningkatkan pemahaman konteks kata, kekurangannya termasuk kompleksitas pemrosesan dan kebutuhan data besar untuk pelatihan yang efektif[10]. *Word2Vec* terdiri dari dua model utama, yaitu *Skip-Gram* dan CBOW (*Continuous Bag of Words*). Pada penelitian lain [11], *Skip-Gram* dipilih karena lebih sesuai dengan dataset yang relatif kecil, yang terdiri dari 2.979 teks dengan ruang vektor (128d, 200d, 300d) dan ukuran jendela (*window size*) 5. Selain itu, dalam penelitian [12], teknik *Word2Vec* digunakan untuk analisis korelasi dan klasifikasi teks dalam kasus kecelakaan kimia, mengeksplorasi hubungan semantik antar kata dalam laporan kecelakaan. Meskipun *Word2Vec* menunjukkan hasil yang baik dalam penelitian ini, performanya sedikit lebih rendah dibandingkan dengan kombinasi *GloVe* + LSTM.

Penelitian ini berkontribusi secara unik dengan mengkhhususkan diri pada analisis sentimen publik terhadap Kurikulum Merdeka, yang masih jarang dieksplorasi dalam literatur. Selain itu, penelitian ini memberikan wawasan baru tentang keefektifan dua model *Word2Vec* dalam konteks pendidikan di Indonesia Berdasarkan karakteristik dataset yang memiliki makna kata yang kompleks, penelitian ini berhipotesis bahwa *model Skip-Gram* lebih efektif dalam memahami dan mengklasifikasikan sentimen yang memiliki nuansa atau detail yang rumit, sementara CBOW mungkin lebih unggul dalam menangkap makna keseluruhan atau konteks umum dari sebuah teks. Diharapkan bahwa temuan dari penelitian ini akan menunjukkan model yang lebih efektif dalam klasifikasi sentimen publik terhadap Kurikulum Merdeka, yang dapat memberikan dasar untuk pengambilan keputusan lebih lanjut dalam pengembangan kurikulum

Tujuan penelitian ini adalah untuk menerapkan algoritma LSTM dalam analisis sentimen terhadap Kurikulum Merdeka, dengan fokus pada perbandingan dua model *Word2Vec*, yaitu *Skip-Gram* dan CBOW. Penelitian ini menggunakan data yang dikumpulkan melalui *crawling* dari Twitter dan YouTube untuk mengklasifikasikan sentimen publik terhadap Kurikulum Merdeka. Berdasarkan penelitian sebelumnya yang telah menunjukkan efektivitas LSTM dan *Word2Vec* dalam memahami konteks kata dan klasifikasi sentimen, studi ini bertujuan untuk mengevaluasi model mana di antara *Skip-Gram* dan CBOW yang paling optimal dalam menangani sentimen terhadap Kurikulum Merdeka.

II. METODE PENELITIAN

Penelitian ini menggunakan data yang diperoleh dari platform media sosial, yaitu YouTube dan Twitter, dengan memanfaatkan *emulator* Google Colab. Tahapan penelitian ini divisualisasikan dalam bentuk diagram alir pada Gambar 1.



Gambar. 1. Diagram Alur Penelitian

Dalam Gambar 1, terlihat bahwa gambar tersebut memberikan gambaran umum tentang alur penelitian. Tahapan tersebut menggambarkan alur penelitian yang sistematis, yang digunakan untuk memastikan setiap langkah penelitian dapat dilakukan secara terstruktur dan efisien.

A. Pengumpulan Data

Penelitian ini memperoleh data melalui proses *crawling* dengan memanfaatkan API dari Twitter dan YouTube. Pada YouTube, data yang dikumpulkan berasal dari ulasan video yang membahas topik "Kurikulum Merdeka." Data dari Twitter dan YouTube dikumpulkan menggunakan dua kata kunci utama, yaitu "kurikulum merdeka" dan "kurikulum baru." Hasil dari proses ini menghasilkan 1.312 data untuk Twitter dan 28.706 untuk YouTube. Hasil setiap *crawling* disimpan dalam format CSV.

Twitter dipilih karena perannya sebagai *platform* utama untuk berbagi informasi dan mengekspresikan perasaan publik [13]. YouTube dipilih karena popularitasnya yang tinggi dan fitur sosialnya, seperti voting dan komentar, yang memungkinkan evaluasi opini publik. Ulasan di YouTube sering kali lebih panjang dan mendalam, memberikan wawasan yang lebih kaya dalam analisis sentimen [14].

B. Text Pre-processing

Text preprocessing dalam penelitian ini meliputi tahapan *Cleaning*, *Case Folding*, *Slangwords Conversion*, *Stopword Removal*, *Tokenisasi*, dan *Stemming*. Tahapan-tahapan ini diterapkan secara berurutan untuk memastikan data yang digunakan siap untuk analisis lebih lanjut.

- 1) *Cleaning* merupakan teknik untuk menghapus karakter khusus, URL, dan simbol yang tidak relevan. Pada penelitian ini, juga dilakukan proses: menghapus baris yang jumlah katanya kurang dari enam, menghapus baris duplikat, menghapus baris dengan nilai *NaN*, dan menghapus kolom yang tidak relevan.
- 2) *CaseFolding* salah satu teknik yang dapat mengubah keseluruhan huruf dalam teks menjadi huruf kecil untuk menjaga konsistensi.

- 3) *Stopwords* merupakan teknik untuk menghapus kata penghubung yang tidak memberikan kontribusi signifikan terhadap makna teks.
- 4) Tokenisasi merupakan teknik untuk memisahkan teks menjadi token atau kata-kata satuan guna mempermudah analisis lebih lanjut.
- 5) *Stemming* merupakan teknik untuk mengubah kata-kata ke bentuk dasarnya

C. Pelabelan Data

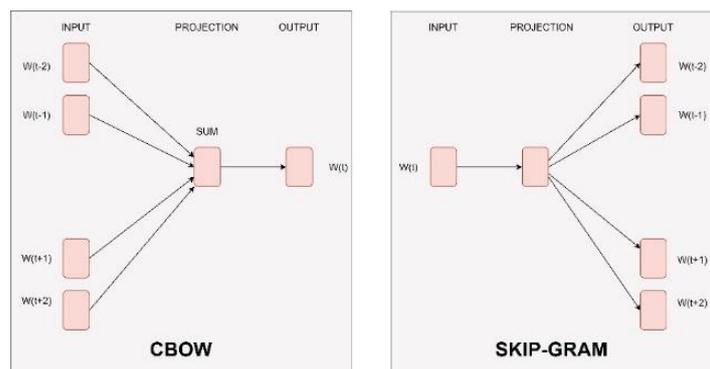
Setelah kedua dataset YouTube dan Twitter digabungkan dan menjalani proses *text preprocessing*, langkah selanjutnya adalah melakukan pelabelan data menggunakan pendekatan berbasis leksikon, yang sangat efisien dalam analisis sentimen pada data besar, terutama dalam konteks media sosial seperti YouTube dan Twitter [15]. Metode ini mengelompokkan data ke dalam kategori sentimen yang relevan, yaitu negatif, positif, dan netral. Hasil dari proses pelabelan kemudian disimpan dalam format CSV, memudahkan penyimpanan dan aksesibilitas. File CSV yang telah dilabeli ini akan digunakan sebagai dataset untuk melatih model yang akan diterapkan dalam penelitian ini.

D. Visualisasi Data Menggunakan Word Cloud

Word Cloud, yang juga dikenal sebagai *TagCloud*, merupakan visualisasi data teks yang menyoroti frekuensi kemunculan kata-kata dalam sebuah dokumen. Kata-kata yang muncul lebih sering akan ditampilkan dalam ukuran yang lebih besar atau warna yang lebih mencolok, memberikan gambaran cepat mengenai topik atau tema yang dominan dalam teks tersebut [16]. *Word Cloud* sering digunakan dalam analisis teks untuk menyederhanakan pemahaman terhadap data yang besar dan kompleks dengan cara yang menarik secara visual.

E. Word2Vec

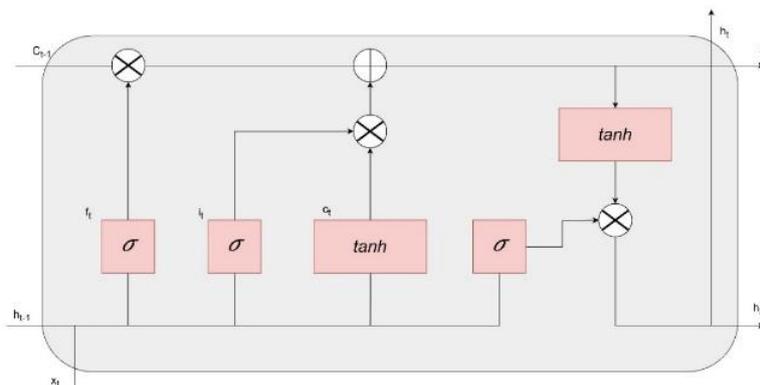
Model Word2Vec digunakan untuk menghasilkan representasi vektor dari kata-kata dalam sebuah dataset, di mana setiap kata direpresentasikan dalam ruang vektor yang sama. Representasi ini memungkinkan pengukuran jarak antara vektor kata untuk menilai kesamaan konteksnya, keberhasilan Word2Vec didorong oleh dua arsitektur utama: *Skip-Gram* dan CBOW. [17]. *Skip-Gram* efektif ketika data pelatihan terbatas, karena mampu menangkap kata atau frasa yang jarang muncul dengan baik. Sebaliknya, sedangkan CBOW mempelajari kata-kata dengan lebih cepat dan akurasi yang tinggi, terutama untuk kata yang kerap sekali digunakan, sehingga lebih efisien dalam merepresentasikan kata kunci umum [18]. Dimensi vektor yang lebih besar dapat menangkap informasi semantik yang lebih kaya, namun ukuran 100 dianggap cukup untuk banyak aplikasi analisis sentimen. Ukuran jendela 5-10 dianggap optimal untuk menangkap konteks yang relevan tanpa menyebabkan pelatihan menjadi terlalu lambat atau kompleks [19]. Pada Gambar 2 ditampilkan arsitektur dari *Skip-Gram* dan CBOW.



Gambar. 2. Arsitektur Skip-Gram dan CBOW

F. Long Short-Term Memory (LSTM)

Algoritma LSTM adalah jenis khusus dari jaringan saraf yang dirancang untuk mengatasi tantangan dalam mempelajari ketergantungan jangka panjang dalam data deret waktu. Sebagai prediktor, jaringan saraf LSTM sangat efektif dalam memahami pola dalam data yang memiliki urutan, baik untuk jangka panjang maupun pendek [7]. LSTM memiliki beberapa lapisan yaitu [20] *input layer*, *forget gate*, *hidden layer* dan *output layer*. Arsitektur LSTM ditunjukkan pada Gambar 3



Gambar. 3. Arsitektur LSTM

Berikut adalah persamaan dari setiap lapisan LSTM:

1) *Input Gate*

$$i_t = \sigma(x_t \cdot w_{ix} + h_{t-1} \cdot w_{ih} + b_i) \quad (1)$$

2) *Forget Gate*

$$f_t = \sigma(x_t \cdot w_{fx} + h_{t-1} \cdot w_{fh} + b_f) \quad (2)$$

3) *Output Gate*

$$o_t = \sigma(x_t \cdot w_{ox} + h_{t-1} \cdot w_{oh} + b_o) \quad (3)$$

4) *Hidden Layer*

$$h_t = o_t \cdot \tanh(c_t) \quad (4)$$

Model LSTM pada penelitian ini dioptimalkan dengan dua *optimizer*, Adam dan RMSprop. Adam dipilih karena kemampuannya menyesuaikan *learning rate* secara dinamis [21], sementara RMSprop unggul dalam menangani gradien yang hilang pada data deret waktu panjang [22]. Jumlah *epoch* ditetapkan pada 10 dengan *batch size* 64. Dua *learning rate*, 0,001 dan 0,0001, diuji untuk melihat respons model. Fungsi aktivasi *Softmax* digunakan untuk mengubah output menjadi probabilitas [23].

G. *Evaluasi*

valuasi model dilakukan dengan menganalisis grafik akurasi dan *loss* selama pelatihan, serta menggunakan matriks kebingungan (*confusion matrix*). Matriks kebingungan merupakan teknik yang efisien untuk menilai performa model klasifikasi dengan menggunakan *dataset* uji, di mana nilai sebenarnya dari kelas sudah diketahui [24]. Grafik akurasi menunjukkan seberapa baik model memprediksi dengan benar, sementara grafik *loss* mengindikasikan tingkat kesalahan model[25]. Adapun *Table confusion matrix* dapat dilihat pada *Table I*.

TABEL I
 CONFUSION MATRIX

	Actual (+)	Actual (-)
Prediksi Class (+)	True Positive(TP)	False Positive(FP)
Prediksi Class (-)	False Negative(FN)	True Negative(TN)

Pengukuran pada *confusion matrix* dapat dievaluasi melalui nilai akurasi, yang dihitung dengan persamaan dibawah :

$$\frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

III. HASIL PEMBAHASAN

Analisis sentimen dalam penelitian ini dilakukan melalui beberapa tahapan dan skenario uji coba, yang mencakup pengumpulan dataset, pemrosesan teks, pelabelan data, dan pengujian model. Parameter uji coba disajikan dalam *Tabel II*.

TABEL II SKENARIO UJI COBA

Skip-Gram dan CBOW	
Dimensi Vector	100
Window	5
LSTM	
Optimizer	Adam, Rmsprop
Learning Rate	0.001, 0.0001

<i>Epochs</i>	10
<i>Batch Size</i>	64
<i>Output Layer</i>	Softmax

A. Dataset

Dataset diperoleh melalui proses *crawling* pada Twitter dan YouTube dengan menggunakan API ID. Dataset mentah yang diperoleh terdiri dari 1.312 data dari Twitter dan 28.706 data dari YouTube. Contoh data dari Twitter dan YouTube yang diperoleh dari hasil *crawling* ditunjukkan pada Tabel III.

TABEL III
 CONTOH DATA TWITTER DAN YOUTUBE CRAWLING

No	Twitter
1	Kali ini saya akan berbagi cerita... Mengenai penerapan Projek P-5 Kurikulum Merdeka yaitu Gaya hidup Berkelanjutan Putik Cantik #ceritakurikulummerdeka #merdekabelajar #festivalkurikulummerdeka2024 #gurudikdas https://t.co/s0e7F3Yd8e
1.313.	@ARSIPAJA Keliatan kayaknya beda kerudung hitam produk kurikulum baru (merdeka?) & yg belakangnya udah agak berumir pasti didikan kurikulum lama.

No	Youtube
1.	Tidak sesuai denga khasanah Infonesia.
28.707.	Tapi emang kurikulum pendidikan kita harus ada perubahan sih, masa potensi anak cuma diukur dari hasil UN dan angka2, padahal itu belum tentu mengcover sepenuhnya kemampuan belajar anak.

Data yang berhasil dikumpulkan disimpan dalam format .csv, dengan total data asli yang digabungkan antara Twitter dan YouTube mencapai 30.018.

B. Text Pre-processing

Proses *pre-processing* ini bertujuan untuk meningkatkan kinerja model dengan beberapa langkah, termasuk pembersihan data (*Cleaning*), perubahan huruf menjadi huruf kecil (*Case Folding*), penghapusan kata tidak penting (*Stopword Removal*), tokenisasi, dan stemming. Setelah melalui proses ini, jumlah dataset berkurang dari 30.018 menjadi 28.706 data. contoh hasil disajikan pada Tabel IV.

TABEL IV
 CONTOH HASIL TEXT PREPROCESSING

Tahapan	Hasil
Kalimat Asli	Betul bossgonta ganti kurikulum -- tidak waras. Karena duit nya bisa di KORUPSI
Cleaning	Betul boss gonta ganti kurikulum tidak waras Karena duit nya bisa di KORUPSI
Case Folding	betul boss gonta ganti kurikulum tidak waras karena duit nya bisa di korupsi
Stopword Removal	betul bos gonta ganti kurikulum waras duit nya di korupsi
Tokenisasi	['betul', 'bos', 'gonta', 'ganti', 'kurikulum', 'waras', 'duit', 'nya', 'di', 'korupsi']
Stemming	betul gonta ganti kurikulum waras duit nya di korupsi

Dalam penelitian ini, penghapusan *stopwords* dan *stemming* dilakukan menggunakan pustaka Sastrawi, yang merupakan pustaka NLP untuk bahasa Indonesia. Penghapusan *stopwords* bertujuan untuk menghapus kata umum yang tidak berkontribusi berarti dalam analisis teks, seperti "dan", "yang", sehingga mengurangi gangguan dalam data. *Stemming* mengubah kata-kata yang terinfleksi ke bentuk dasarnya, seperti "bermain" menjadi "main" dan "mengajar" menjadi "ajar," guna memastikan konsistensi dan meningkatkan akurasi dalam analisis teks.

C. Pelabelan Data

Pelabelan menggunakan pendekatan *lexicon*, di mana pendekatan ini menggunakan kamus sebagai referensi bahasa. Tahapannya mencakup tiga proses utama: mengimpor *lexicon* untuk sentimen positif, netral, dan negatif; memecah teks menjadi kata-kata untuk mempermudah evaluasi skor sentimen; serta menghitung skor dengan

E. Penerapan Model

Setelah pelabelan dilakukan pada 28.706 data, langkah berikutnya adalah menerapkan model *Word2Vec* dan LSTM. Dataset dibagi menjadi rasio 80:20, dengan 80% untuk melatih model, sementara 20% digunakan untuk menguji model. Penelitian ini membandingkan hasil *Skip-Gram* dan CBOV dengan algoritma LSTM. Hasil pengujian tersebut dibagi menjadi beberapa skenario. Tabel VII menunjukkan hasil uji coba LSTM + *Word2Vec* dengan arsitektur *Skip-Gram*. Tabel VIII menunjukkan hasil uji coba LSTM + *Word2Vec* dengan arsitektur CBOV.

TABEL VII
 HASIL PENGUJIAN SKIP-GRAM

Skenario	Optimizer	Learning Rate	Epoch	Batch Size	Window	Dimensi Vector	Accuracy
1	Adam	0,001	10	64	5	100	0,97
2	Adam	0,0001	10	64	5	100	0,93
3	RMSProp	0,001	10	64	5	100	0,97
4	RMSProp	0,0001	10	64	5	100	0,87

TABEL VIII
 HASIL PENGUJIAN CBOV

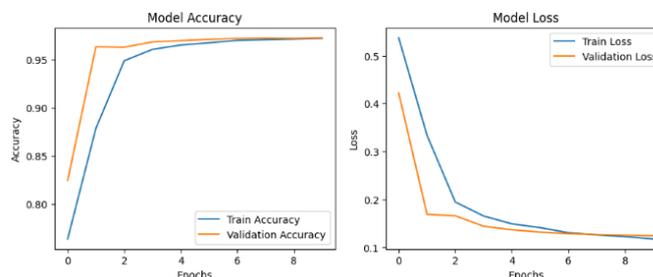
Skenario	Optimizer	Learning Rate	Epoch	Batch Size	Window	Dimensi Vector	Accuracy
1	Adam	0,001	10	64	5	100	0,97
2	Adam	0,0001	10	64	5	100	0,96
3	RMSProp	0,001	10	64	5	100	0,97
4	RMSProp	0,0001	10	64	5	100	0,96

F. Evaluasi

Performa model selama proses dapat dilihat dari beberapa pengujian yang dilakukan menggunakan *Skip-Gram* dan CBOV menggunakan algoritma LSTM.

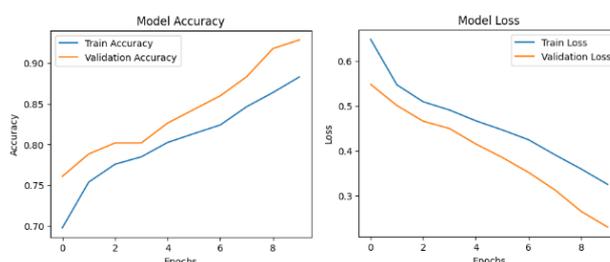
1) *Skip-Gram*

Pada skenario pertama uji coba dengan *Skip-Gram*, dapat dilihat pada Gambar 7.



Gambar. 7. Skenario 1 *Skip-Gram*

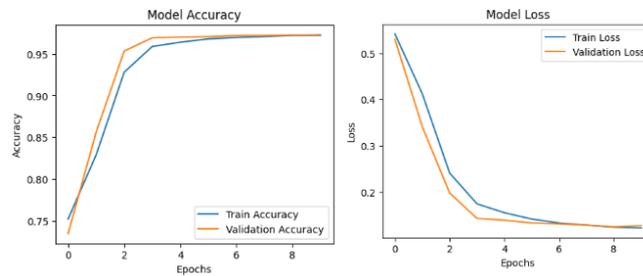
Pada skenario 1 dengan parameter pengujian pada Tabel 7 dapat dilihat bahwa pada skenario 1, model menggunakan *optimizer Adam* dengan *learning rate* 0.001. Model mencapai akurasi tinggi sebesar 0,97, dengan peningkatan akurasi yang cepat dan stabil selama pelatihan. Grafik menunjukkan penurunan *loss* yang signifikan, mengindikasikan bahwa model berhasil belajar dengan baik dan efektif. Uji coba pada Skenario 2 pada *Skip-Gram* disajikan pada Gambar 8.



Gambar. 8. Skenario 2 *Skip-Gram*

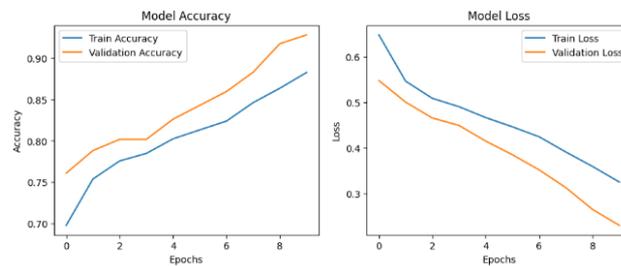
Pada skenario 2 yang ditunjukkan pada Gambar 8, model tetap menggunakan *optimizer Adam*, tetapi dengan *learning rate* yang lebih kecil, yaitu 0,0001. Akurasi akhir yang dicapai adalah 0,93. Grafik menunjukkan bahwa model mengalami peningkatan akurasi dan penurunan *loss* secara konsisten, namun dengan

kecepatan yang lebih lambat dibandingkan skenario 1, karena learning rate yang lebih kecil membuat proses pembelajaran lebih lambat. Selanjutnya pada skenario 3, yang dapat dilihat pada Gambar 9.



Gambar. 9. Skenario 3 Skip-Gram

Skenario 3 menggunakan *optimizer* RMSprop dengan learning rate 0,001. Model berhasil mencapai akurasi tinggi sebesar 0,97, mirip dengan hasil skenario 1. Grafik menunjukkan bahwa model belajar dengan sangat baik, dengan peningkatan akurasi yang cepat dan penurunan *loss* yang signifikan selama beberapa *epoch* pertama, menunjukkan kinerja yang optimal dengan kombinasi parameter ini. Skenario 4 dapat dilihat pada Gambar 10.

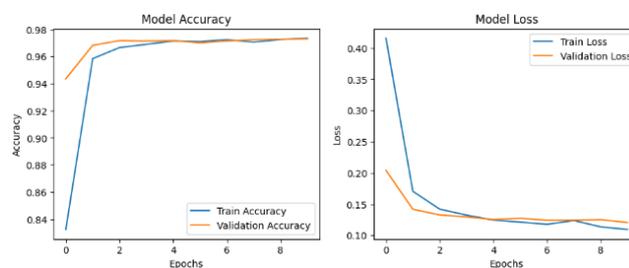


Gambar. 10. Skenario 4 Skip-Gram

Pada skenario 4, model menggunakan *optimizer* RMSprop dengan *learning rate* 0,0001. Akurasi akhir yang dicapai adalah 0,86, yang lebih rendah dibandingkan skenario sebelumnya. Grafik menunjukkan bahwa model mengalami peningkatan akurasi dan penurunan *loss* secara stabil, namun hasil akhir lebih rendah karena *learning rate* yang lebih kecil membuat model belajar lebih lambat dan mungkin belum mencapai potensi performa optimal dalam jangka waktu yang diberikan.

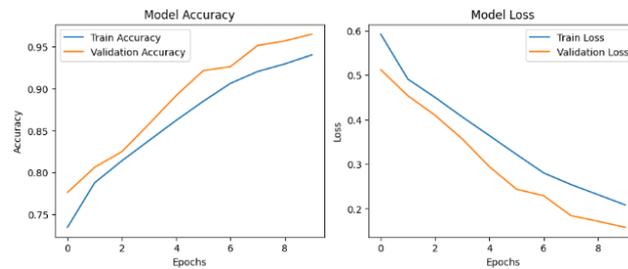
2) CBOW

Pada scenario pertama uji coba dengan CBOW, dapat dilihat pada Gambar 11.



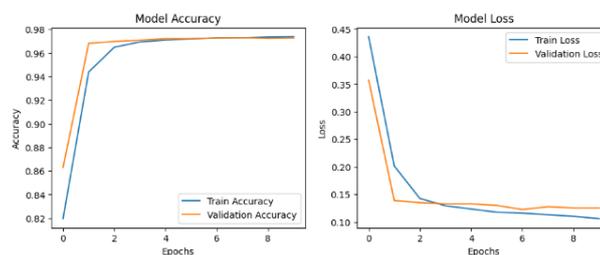
Gambar. 11. Skenario 1 CBOW

Pada skenario 1 dengan CBOW, model menggunakan *optimizer* Adam dengan *learning rate* 0,001. Model mencapai akurasi tinggi sebesar 0,97, dengan tren akurasi yang cepat dan stabil selama pelatihan. Grafik *loss* menunjukkan penurunan yang signifikan, yang mengindikasikan bahwa model berhasil belajar dengan baik dan efektif dalam meminimalkan kesalahan prediksi. Selanjutnya pada skenario 2 CBOW, dapat dilihat pada Gambar 12.



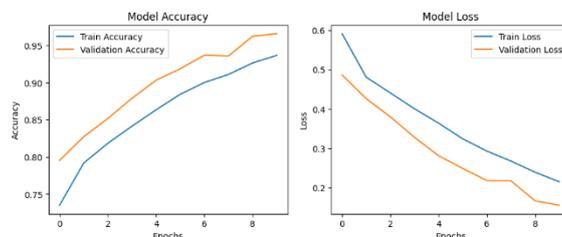
Gambar. 12. Skenario 2 CBOW

Pada skenario 2, model tetap menggunakan *optimizer Adam* tetapi dengan *learning rate* yang lebih kecil, yaitu 0,0001. Akurasi akhir yang dicapai adalah 0,96. Grafik menunjukkan bahwa meskipun ada peningkatan akurasi dan penurunan *loss* yang konsisten, kecepatan peningkatan sedikit lebih lambat dibandingkan dengan skenario pertama karena *learning rate* yang lebih kecil. Skenario 3 CBOW selanjutnya dapat dilihat pada Gambar 13.



Gambar. 13. Skenario 3 CBOW

Skenario 3 menggunakan *optimizer RMSprop* dengan *learning rate* 0,001. Model berhasil mencapai akurasi tinggi sebesar 0,97, dengan grafik yang menunjukkan peningkatan akurasi yang cepat dan penurunan *loss* yang signifikan selama beberapa *epoch* pertama. Kombinasi RMSprop dengan *learning rate* 0,001 memberikan hasil yang optimal dalam skenario CBOW ini. Skenario 4 dapat dilihat pada Gambar 14.



Gambar. 14. Skenario 4 CBOW

Pada skenario 4, model menggunakan *optimizer RMSprop* dengan *learning rate* 0,0001. Akurasi yang dicapai adalah 0,96. Grafik menunjukkan peningkatan akurasi yang stabil dan penurunan *loss* yang konsisten. Meskipun *learning rate* yang digunakan lebih kecil, model tetap mampu mempelajari pola data dengan baik, namun dengan sedikit penurunan performa dibandingkan skenario dengan *learning rate* yang lebih tinggi

Berdasarkan hasil pengujian, kombinasi *optimizer Adam* dengan *learning rate* 0,001 terbukti memberikan performa terbaik pada dataset yang digunakan, karena mampu mencapai konvergensi dengan cepat dan stabil pada data yang kompleks dan bervariasi. Kecepatan dan *fleksibilitas optimizer Adam* dalam menyesuaikan *learning rate* selama proses pelatihan sangat mendukung model dalam mengelola kompleksitas data dari media sosial, yang memiliki variasi tinggi dalam konten teks dan struktur kalimat. Sebaliknya, penggunaan *learning rate* yang lebih rendah (0,0001), meskipun menghasilkan pembelajaran yang lebih halus, cenderung memperlambat konvergensi, terutama pada arsitektur *Skip-Gram*. Hal ini menunjukkan bahwa dataset Kurikulum Merdeka memerlukan kecepatan pembelajaran yang lebih tinggi untuk mencapai akurasi optimal, terutama dalam menangani variasi dan ketidakpastian dalam data.

Hasil ini mengindikasikan bahwa model *deep learning* seperti LSTM lebih unggul dalam menangani kompleksitas data teks dibandingkan metode tradisional seperti *Naïve Bayes*[26] dan SVM[27] dalam

analisis sentimen yang sama terkait Kurikulum Merdeka. Penelitian sebelumnya yang menggunakan Naïve Bayes mencapai akurasi sebesar 80,20%, sementara SVM mencapai 91,82%. Penelitian ini berhasil mencapai hasil yang signifikan dalam analisis sentimen terhadap Kurikulum Merdeka, dengan akurasi mencapai 97%, melampaui studi sebelumnya yang menggunakan kombinasi Word2Vec dan LSTM untuk analisis sentiment pada ulasan hotel berbahasa Indonesia, yang hanya mencapai akurasi sebesar 85,96% [8].

IV. KESIMPULAN

Hasil uji coba menunjukkan bahwa kedua arsitektur *Word2Vec*, yaitu *Skip-Gram* dan CBOW, mampu memberikan performa yang baik dalam analisis sentimen Kurikulum Merdeka menggunakan algoritma LSTM. CBOW unggul dengan akurasi mencapai hingga 97% dengan stabilitas yang lebih tinggi dibandingkan *Skip-Gram*, meskipun *Skip-Gram* juga menunjukkan hasil yang kuat. Penggunaan optimizer Adam dan RMSprop pada kedua arsitektur memberikan hasil optimal, dengan Adam sedikit lebih unggul dalam hal stabilitas dan kecepatan konvergensi. *Learning rate* yang lebih kecil (0,0001) menghasilkan pembelajaran yang lebih lambat, sementara *learning rate* yang lebih tinggi (0,001) memberikan akurasi yang sedikit lebih baik. Dataset Kurikulum Merdeka ini terdiri dari 7.170 data positif, 189 data negatif, dan 16.217 data netral. Hasil penelitian ini menunjukkan bahwa kombinasi *Word2Vec* dan LSTM efektif untuk analisis sentimen, dengan CBOW menjadi arsitektur yang paling optimal dalam skenario ini.

DAFTAR PUSTAKA

- [1] M. P. Prof. Dr. H. E. Mulyasa, *IMPLEMENTASI KURIKULUM MERDEKA*. Jakarta Timur: PT.Bumi Aksara, 2023. [Online]. Available: https://books.google.co.id/books?id=ec_hEAAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- [2] W. Darmawan, M. Kurniawan Faizal, W. Setianto, and W. Hapsoro, "Analisis Sentimen Penerapan Kurikulum Merdeka Pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor Dengan Forward Selection," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 12, no. 1, 2023, doi: 10.30591/smartcomp.v12i1.4634.
- [3] M. Wankhade, A. C. S. Rao, and C. Kulkarni, *A survey on sentiment analysis methods, applications, and challenges*, vol. 55, no. 7. Springer Netherlands, 2022. doi: 10.1007/s10462-022-10144-1.
- [4] J. F. Sánchez-Rada and C. A. Iglesias, "Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison," *Inf. Fusion*, vol. 52, no. December 2018, pp. 344–356, 2019, doi: 10.1016/j.inffus.2019.05.003.
- [5] R. P. Pratama and A. Tjahyanto, "The influence of fake accounts on sentiment analysis related to COVID-19 in Indonesia," *Procedia Comput. Sci.*, vol. 197, no. 2021, pp. 143–150, 2021, doi: 10.1016/j.procs.2021.12.128.
- [6] N. Parveen, P. Chakrabarti, B. T. Hung, and A. Shaik, "Twitter sentiment analysis using hybrid gated attention recurrent network," *J. Big Data*, 2023, doi: 10.1186/s40537-023-00726-3.
- [7] R. Bolboacă and P. Haller, "Performance Analysis of Long Short-Term Memory Predictive Neural Networks on Time Series Data," *Mathematics*, vol. 11, no. 6, 2023, doi: 10.3390/math11061432.
- [8] P. F. Muhammad, R. Kusumaningrum, and A. Wibowo, "ScienceDirect ScienceDirect Sentiment Analysis Using Word2vec And Long Short-Term Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews Memory (LSTM) For Indonesian Hotel Reviews," *Procedia Comput. Sci.*, vol. 179, no. 2020, pp. 728–735, 2021, doi: 10.1016/j.procs.2021.01.061.
- [9] P. Rakshit and A. Sarkar, "A supervised deep learning-based sentiment analysis by the implementation of Word2Vec and GloVe Embedding techniques," 2024, doi: 10.1007/s11042-024-19045-7.
- [10] H. Imaduddin, L. A. Kusumaningtyas, and F. Y. A'la, "Application of LSTM and GloVe Word Embedding for Hate Speech Detection in Indonesian Twitter Data," *Ing. des Syst. d'Information*, vol. 28, no. 4, pp. 1107–1112, 2023, doi: 10.18280/isi.280430.
- [11] N. R. Bhowmik, M. Arifuzzaman, and M. R. H. Mondal, "Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms," *Array*, vol. 13, p. 100123, 2022, doi: 10.1016/j.array.2021.100123.
- [12] S. Jing *et al.*, "Correlation analysis and text classification of chemical accident cases based on word embedding," *Process Saf. Environ. Prot.*, vol. 158, no. December, pp. 698–710, 2022, doi: 10.1016/j.psep.2021.12.038.
- [13] M. Mahdikhani, "Predicting the popularity of tweets by analyzing public opinion and emotions in different stages of Covid-19 pandemic," *Int. J. Inf. Manag. Data Insights*, vol. 2, no. 1, p. 100053, 2022, doi: 10.1016/j.jjimei.2021.100053.
- [14] R. Giri, M. Sirsath, and H. T. Kanakia, "Youtube Comments Sentiment Analysis," *2024 IEEE 9th Int. Conf. Conver. Technol. I2CT 2024*, no. May, 2024, doi: 10.1109/I2CT61223.2024.10544022.
- [15] C. Kaushik and A. Mishra, "A Scalable, Lexicon Based Technique for Sentiment Analysis," *Int. J. Found. Comput. Sci. Technol.*, vol. 4, no. 5, pp. 35–56, 2014, doi: 10.5121/ijfst.2014.4504.
- [16] M. G. Moldovan, D. C. Dabija, L. Stanca, and C. B. Pocol, "A Qualitative Study on the Consumer Behaviour Related to Food Waste: Romanian Perspectives through Word Cloud and Sentiment Analysis," *Sustain.*, vol. 16, no. 10, 2024, doi: 10.3390/su16104193.
- [17] C. O. Truică and E. S. Apostol, "It's All in the Embedding! Fake News Detection Using Document Embeddings," *Mathematics*, vol. 11, no. 3, 2023, doi: 10.3390/math11030508.
- [18] D. K. Sharma, B. Singh, S. Agarwal, N. Pachauri, A. A. Alhussan, and H. A. Abdallah, "Sarcasm Detection over Social Media Platforms Using Hybrid Ensemble Model with Fuzzy Logic," *Electron.*, vol. 12, no. 4, pp. 1–21, 2023, doi: 10.3390/electronics12040937.
- [19] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *1st Int. Conf. Learn. Represent. ICLR 2013 - Work. Track Proc.*, pp. 1–12, 2013.
- [20] L. S. Memory and S. Power, "Efficient Energy Management Based on Convolutional," 2021.
- [21] L. Luo, Y. Xiong, Y. Liu, and X. Sun, "Adaptive gradient methods with dynamic bound of learning rate," *7th Int. Conf. Learn. Represent. ICLR 2019*, no. 2018, pp. 1–19, 2019.
- [22] S. M. Zaman, M. M. Hasan, R. I. Sakline, D. Das, and M. A. Alam, "A Comparative Analysis of Optimizers in Recurrent Neural Networks for Text Classification," *2021 IEEE Asia-Pacific Conf. Comput. Sci. Data Eng. CSDE 2021*, pp. 1–6, 2021, doi: 10.1109/CSDE53843.2021.9718394.
- [23] X. Li, X. Li, D. Pan, and D. Zhu, "On the learning property of logistic and softmax losses for deep neural networks," *AAAI 2020 - 34th AAAI Conf. Artif. Intell.*, no. 9, pp. 4739–4746, 2020, doi: 10.1609/aaai.v34i04.5907.
- [24] N. E. Ramli, Z. R. Yahya, and N. A. Said, "Confusion Matrix as Performance Measure for Corner Detectors," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 29, no. 1, pp. 256–265, 2022, doi: 10.37934/araset.29.1.256265.
- [25] K. Tutuncu, I. Cinar, R. Kursun, and M. Koklu, "Edible and Poisonous Mushrooms Classification by Machine Learning Algorithms," *2022 11th*

- Mediterr. Conf. Embed. Comput. MECO 2022*, no. June, 2022, doi: 10.1109/MECO55406.2022.9797212.
- [26] A. Ariefah, Widodo, and M. Nugraheni, "Sentiment Analysis for Curriculum of Independent Learning Based on Naïve Bayes with Laplace Estimator," *2023 Int. Conf. Inf. Technol. Res. Innov. ICITRI 2023*, pp. 157–161, 2023, doi: 10.1109/ICITRI59340.2023.10249320.
- [27] I. Ramadhanti, Widodo, and M. Nugraheni, "Sentiment Analysis of Public Opinion on Twitter about the Implementation of the Merdeka Curriculum Using the Support Vector Machine Algorithm," *Proc. ICMERALDA 2023 - Int. Conf. Model. E-Information Res. Artif. Learn. Digit. Appl.*, pp. 86–91, 2023, doi: 10.1109/ICMERALDA60125.2023.10458193.