



IMPLEMENTASI BI-LSTM DENGAN EKSTRAKSI FITUR WORD2VEC UNTUK PENGEMBANGAN ANALISIS SENTIMEN APLIKASI IDENTITAS KEPENDUDUKAN DIGITAL

Romario Onsu¹, Daniel Febrian Sengkey², Feisy Diane Kambey³

^{1, 2, 3}Teknik Elektro, Universitas Sam Ratulangi
Manado, Sulawesi Utara, Indonesia 95115

romarioonsu026@student.unsrat.ac.id, danielsengkey@unsrat.ac.id, feisykambey@unsrat.ac.id

Abstract

The Indonesian government is striving to enhance digital public services, including the Digital Identity Application (IKD) launched in 2022 by the Directorate General of Population and Civil Registration. Since its launch, IKD has received various responses from the public. User reviews on Google Play Store indicate a decline in ratings from June to December 2023. Review analysis is essential to understand user satisfaction, identify issues, and guide application improvements. This study aims to perform sentiment analysis on IKD user reviews using Bidirectional Long Short-Term Memory (Bi-LSTM) and Word2Vec methods. Bi-LSTM and Word2Vec are used to develop sentiment analysis from previous research that still used Machine Learning methods. This research is expected to contribute to the development of sentiment analysis models using Deep Learning for the IKD application. Review data was collected from the Google Play Store using scraping techniques for the period January-December 2023 and categorized into positive and negative. The Bi-LSTM model was trained with Word2Vec CBOW and Skip-Gram variations with dimensions of 100, 200, and 300. The results show that the combination of Bi-LSTM and Word2Vec CBOW with a dimension of 200 and a data split ratio of 80/20 produced the highest accuracy of 96.06%, with a precision of 96.44%, recall of 95.64%, and an f1-score of 96.04%. All combinations of Bi-LSTM and Word2Vec outperformed other Machine Learning algorithms.

Keywords: Bi-LSTM, IKD, Sentiment Analysis, Text Mining, Word2Vec

Abstrak

Pemerintah Indonesia berupaya meningkatkan layanan publik berbasis digital, termasuk aplikasi Identitas Kependudukan Digital (IKD) yang diluncurkan pada 2022 oleh Dirjen Kependudukan dan Pencatatan Sipil. Sejak diluncurkan, IKD mendapat berbagai tanggapan dari masyarakat. Data ulasan di Google Play Store menunjukkan penurunan rating dari Juni hingga Desember 2023. Analisis ulasan penting untuk memahami kepuasan pengguna dan mengidentifikasi masalah serta memandu perbaikan aplikasi. Penelitian ini bertujuan melakukan analisis sentimen ulasan pengguna IKD menggunakan metode Bidirectional Long Short-Term Memory (Bi-LSTM) dan Word2Vec. Bi-LSTM dan Word2Vec digunakan untuk mengembangkan analisis sentimen dari penelitian sebelumnya yang masih menggunakan metode Machine Learning. Penelitian ini diharapkan berkontribusi dalam pengembangan model analisis sentimen menggunakan Deep Learning untuk aplikasi IKD. Data ulasan dikumpulkan dari Google Play Store dengan teknik scraping pada periode Januari-Desember 2023 dan dibagi menjadi kategori positif dan negatif. Model Bi-LSTM dilatih dengan variasi Word2Vec CBOW dan Skip-Gram dengan dimensi 100, 200, dan 300. Hasil penelitian menunjukkan kombinasi Bi-LSTM dan Word2Vec CBOW dengan dimensi 200 dan proporsi data 80/20 menghasilkan akurasi tertinggi sebesar 96,06%, dengan precision 96,44%, recall 95,64%, dan f1 score 96,04%. Semua kombinasi Bi-LSTM dan Word2Vec menunjukkan hasil lebih tinggi dibandingkan algoritma Machine Learning lainnya.

Kata kunci: Analisis Sentimen, Bi-LSTM, IKD, Text Mining, Word2Vec

1. PENDAHULUAN

Dalam beberapa tahun terakhir, pemerintah Indonesia berupaya mengembangkan layanan publik menjadi digital untuk mencapai transformasi digital di era teknologi,

diharapkan bisa memudahkan masyarakat di berbagai sektor, terkhusus dalam layanan publik. Transformasi layanan publik menjadi digital bertujuan meningkatkan transparansi, efisiensi, dan responsivitas terhadap

kebutuhan warga negara[1]. Salah satu upaya mengembangkan layanan publik berbasis digital. Pemerintah Indonesia melalui Direktorat Jenderal Kependudukan dan Pencatatan Sipil meluncurkan aplikasi Identitas Kependudukan Digital (IKD) di tahun 2022. Aplikasi Identitas Kependudukan Digital mengintegrasikan berbagai layanan seperti kesehatan, pelayanan sosial, pendidikan, perpajakan, perbankan, dan lainnya ke dalam satu aplikasi untuk aktivasi[2]. Sejak diluncurkan, IKD telah mendapatkan berbagai tanggapan dari masyarakat, terutama melalui ulasan dari pengguna di *Google Play Store*, yang berisi berbagai macam opini masyarakat berupa pujian, kritik sampai masukan terhadap aplikasi. Aplikasi IKD mengalami penurunan *rating* dalam rentang waktu yang cukup lama dari Juni–Desember 2023. Ulasan pengguna dapat menjadi sumber informasi berharga mengenai tanggapan masyarakat dan umpan balik terhadap aplikasi, serta mengetahui kekurangan dan potensi perbaikan yang dapat diperbaiki. Namun, dengan begitu banyak jumlah ulasan, sulit untuk membedakan secara manual, ulasan yang memiliki sentimen bersifat positif dan bersifat negatif[3].

Dalam menghadapi masalah yang ditemukan, diperlukan analisis sentimen untuk mengklasifikasikan sentimen secara otomatis terhadap ulasan pengguna aplikasi IKD dari *Google Play Store*. Penggunaan analisis sentimen terhadap ulasan pengguna aplikasi layanan digital seperti IKD, Sangat berguna bagi Pemerintah dan pengembang dalam mengukur seberapa baik aplikasi yang di jalankan mampu menyediakan layanan publik digital dengan baik kepada pengguna[4][5]. Dalam upaya untuk meningkatkan kinerja aplikasi yang bisa diterima masyarakat, analisis sentimen terhadap aplikasi Identitas Kependudukan Digital sangat perlu dilakukan sebagai pertimbangan untuk evaluasi perkembangan aplikasi dan menilai potensi perbaikan. Penelitian terdahulu terkait analisis sentimen pada aplikasi Identitas Kependudukan Digital masih menggunakan pengolahan dengan metode *Machine Learning Konvensional*[2][6][7]. Dikarenakan penelitian terkait aplikasi IKD sebelumnya hanya menggunakan metode *Machine Learning* konvensional, terdapat kebutuhan yang belum terpenuhi untuk mengeksplorasi potensi metode *Deep Learning* terhadap analisis sentimen ulasan aplikasi IKD. Model *Deep Learning* seperti *Artificial Neural Networks*, sangat baik dalam memproses volume data yang besar dengan efisien, dengan mengekstraksi pola-pola kompleks dari data teks[8]. Model *Deep Learning* ini, mempelajari banyak lapisan representasi atau fitur dari data, yang menghasilkan hasil prediksi paling akurat[9]. Di beberapa penelitian analisis sentimen yang membandingkan antara penerapan metode *Machine Learning* dan *Deep Learning*. *Deep Learning* memiliki performa lebih baik dibandingkan dengan *Machine Learning*[10][11][12]. Algoritma *Deep Learning* seperti *Recurrent Neural Network* (RNN) dengan turunannya seperti *Long Short-Term Memory* (LSTM) yang banyak digunakan untuk analisis sentimen.

Sebelum pembentukan model *Deep Learning*, data teks yang akan dianalisis sentimen harus diekstraksi fiturnya terlebih dahulu. Dalam proses ini, kata-kata dalam teks akan diubah menjadi vektor. Algoritma *Deep Learning* seperti LSTM cukup efektif jika disandingkan dengan *Word2Vec* untuk ekstraksi fitur pada teks[13][14]. Namun LSTM hanya memproses data *input* melalui satu arah, sehingga model LSTM dikembangkan dengan *input*-nya menjadi dua arah. Memungkinkannya untuk menangkap informasi kontekstual dengan lebih efektif[15]. Metode *Bidirectional LSTM* yang dikombinasikan dengan *Word2Vec* menunjukkan hasil yang sangat baik, dan melebihi LSTM biasa. Penelitian oleh [16], mengenai analisis ulasan destinasi wisata di Pulau Bali, menggunakan *Bidirectional LSTM* mendapatkan hasil yang sangat baik dengan akurasi sebesar 96,86%. Beberapa penelitian yang menggunakan Bi-LSTM dengan *Word2Vec* juga menunjukkan performa yang sangat baik dibandingkan dengan algoritma lain[17][18].

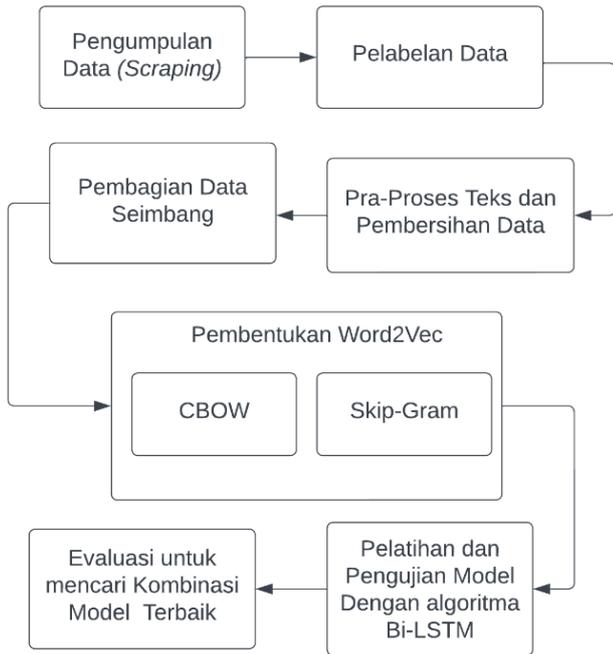
Sebagaimana telah dibahas pada penelitian-penelitian sebelumnya, Bi-LSTM memiliki keunggulan pada pengolahan analisis sentimen terutama dengan *dataset* yang besar. Oleh karena itu, penelitian ini akan menggunakan Bi-LSTM sebagai algoritma *Deep Learning* dengan *Word2Vec* sebagai fitur ekstraksi yang akan di terapkan untuk mengembangkan metode analisis sentimen pada ulasan pengguna aplikasi IKD, yang pada penelitian-penelitian sebelumnya belum menggunakan penerapan metode *Deep Learning*.

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan pengguna aplikasi Identitas Kependudukan Digital menggunakan metode *Bidirectional Long Short-Term Memory* dan *Word2Vec*. Kombinasi metode ini dipilih berdasarkan hasil dari uraian di atas menunjukkan bahwa, Bi-LSTM dan *Word2Vec* merupakan mode yang sangat baik untuk digunakan dalam analisis sentimen. Model analisis sentimen sangat berguna bagi pengembang aplikasi dan Pemerintah untuk mengukur keberhasilan manajemen aplikasi IKD dan merupakan bahan evaluasi untuk mengembangkan IKD. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi sebagai pengembangan dari metode analisis sentimen untuk menganalisis sentimen ulasan masyarakat terhadap aplikasi Identitas Kependudukan Digital menggunakan penerapan metode *Deep Learning*, agar bisa mempermudah dalam menganalisis sentimen publik.

2. METODE PENELITIAN

Penelitian ini melibatkan serangkaian tahapan dan pengujian dalam pelaksanaan proses penelitian. Tahapan yang dilakukan meliputi pengumpulan data, pelabelan data, *text preprocessing*, pembentukan model *Word2Vec*, pelatihan model, pengujian menggunakan Bi-LSTM dengan variasi *Word2Vec*, dan evaluasi model. Evaluasi model dilakukan menggunakan *Confusion Matrix* untuk mencari model terbaik. Metode penelitian ini mencakup langkah-

langkah yang sistematis untuk memastikan setiap tahap dilakukan dengan baik dan menghasilkan model yang optimal. Pada gambar 1 di bawah ini merupakan diagram alir penelitian yang akan menjadi alur dari penelitian.

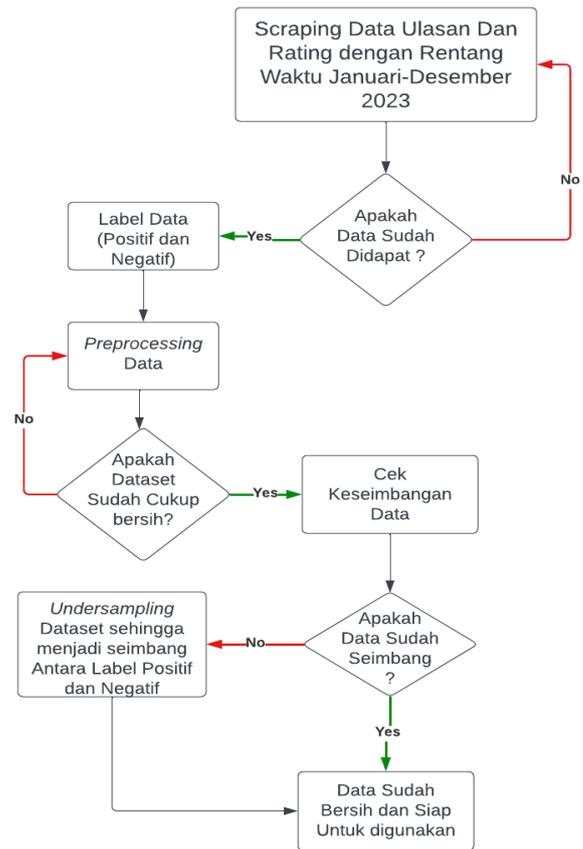


Gambar 1. Diagram Alir Penelitian

Penelitian ini menggunakan data ulasan dan *rating* yang dikumpulkan menggunakan pustaka Python *Google Play Scraper*. Setelah data terkumpul, dilakukan pelabelan dan pembersihan data (*preprocessing*). Setelah data dinilai sudah cukup bersih maka selanjutnya adalah dengan membagi data dengan nilai yang seimbang antara data yang berlabel sentimen positif dan sentimen negatif agar *dataset* yang akan digunakan berkualitas baik untuk digunakan dalam model. Berikutnya, dibuat model Word2Vec untuk memetakan kata-kata dalam ulasan ke dalam vektor numerik yang merepresentasikan makna kata. Model klasifikasi sentimen kemudian dilatih menggunakan algoritma Bi-LSTM dalam menangani data teks berurutan. Kedua variasi Word2Vec diuji untuk menentukan performa model terbaik dengan kombinasi Bi-LSTM. Terakhir, model dievaluasi menggunakan *Confusion Matrix* untuk mengukur akurasi, presisi, *recall*, dan F1-score.

2.1 Pemrosesan Data

Penelitian ini melakukan serangkaian pemrosesan data secara komprehensif seperti pada gambar 2 di bawah ini yang merupakan Diagram Alir Pemrosesan Data.



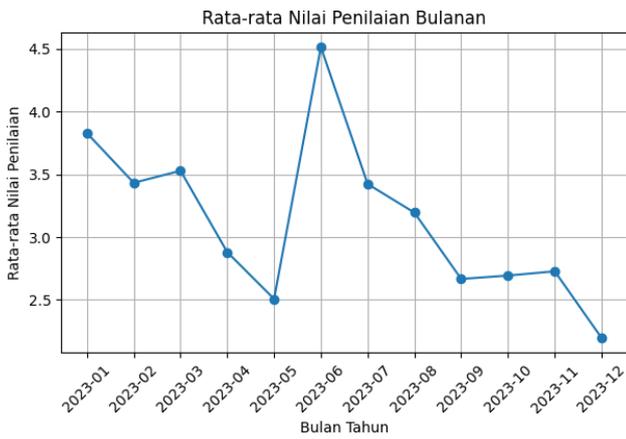
Gambar 2. Diagram Alir Pemrosesan Data

Pada pemrosesan data, pertama dilakukan pengumpulan data lewat *Google Play Store* dengan menggunakan *scraping*. Selanjutnya dilakukan Pelabelan data dengan positif dan negatif. Berikutnya masuk pada tahap *preprocessing*. Setelah data terlihat cukup bersih untuk digunakan, dilakukan penyeimbang data dengan melakukan *undersampling* sehingga data seimbang dan siap untuk digunakan.

2.1.1 Pengumpulan Data

Pada pengumpulan data, penelitian ini melakukan *scraping* data ulasan dan *rating* pada *Google Play Store* terhadap aplikasi Identitas Kependudukan Digital menggunakan pustaka Python *Google-Play-Scraper*. Pengambilan data menggunakan *Google Collab* dengan Bahasa pemrograman Python. Data yang diambil pada rentang waktu 1 tahun yaitu dari bulan Januari 2023–Desember 2023 dengan jumlah 14.000 data ulasan beserta *rating* yang diberikan pada ulasan aplikasi IKD.

Berikut pada gambar 3 di bawah ini merupakan grafik dari rata-rata penilaian terhadap aplikasi IKD dengan rentang waktu 1 tahun antara Januari 2023–Desember 2023.



Gambar 3. Grafik Rata-Rata Penilaian Aplikasi IKD

Pada grafik penilaian aplikasi IKD di atas terlihat pada bulan Juni sampai dengan bulan Desember 2023, mengalami penurunan *rating* terhadap aplikasi IKD, ini menunjukkan adanya penurunan tingkat kepuasan pengguna terhadap aplikasi IKD pada periode rentang waktu tersebut. Berikut contoh data ulasan dan *rating* terhadap aplikasi IKD yang bisa dilihat pada tabel 1.

Tabel 1. Contoh Ulasan dan *Rating* pada *Dataset*

Ulasan	Rating
Pas di upgrade ke android 12 jadi gak bisa kebuka, malah menyambungkan terus	2
Tak perlu repot jg tak ribet	5
Mantap Dan Sukses..	5
scan barcode harus dtg ke dukcapil...wasting time..katanya digital, hadeuuh..	1

2.1.2 Pelabelan Data

Pada tahap pelabelan data dilakukan *labeling* secara manual dengan mengategorikan ulasan yang bernilai 4 dan 5 sebagai Positif sedangkan 1 dan 2 sebagai Negatif. Kemudian Ulasan yang bernilai 3 akan dihapus dari *dataset*. Sehingga data yang digunakan hanya Positif dan Negatif. Pada tabel 2 merupakan contoh data ulasan yang ada pada *dataset*.

Tabel 2. Contoh Data Ulasan dan Pelabelan pada *Dataset*

Ulasan	Sentiment
Pas di upgrade ke android 12 jadi gak bisa kebuka, malah menyambungkan terus	Negatif
Tak perlu repot jg tak ribet	Positif
Mantap Dan Sukses..	Positif
scan barcode harus dtg ke dukcapil...wasting time..katanya digital, hadeuuh..	Negatif

2.1.3 Text Preprocessing

Pra-proses teks atau *text preprocessing* adalah serangkaian proses yang digunakan untuk menyiapkan teks mentah agar lebih siap untuk dianalisis. *Text preprocessing* sangat penting dan diperlukan dalam pengolahan *Natural Language Processing* (NLP)[19]. Tujuan utama *text preprocessing* adalah mengurangi kebisingan dan menyederhanakan teks sehingga analisis lanjutan atau algoritma pembelajaran mesin dapat bekerja lebih efisien dan efektif. Berikut serangkaian tahapan pada *text preprocessing* yang akan digunakan. Pada tahapan *text preprocessing* semua dilakukan dengan menggunakan *library* dari *Natural Language Toolkit* (NLTK), untuk memudahkan proses *text preprocessing*.

a. Case Folding

Case folding adalah salah satu bentuk *teks preprocessing*, untuk mengubah semua huruf dalam dokumen menjadi huruf kecil. Pada tabel 3 di bawah ini merupakan contoh penggunaan *case folding* pada *dataset*.

Tabel 3. Contoh *Case Folding* pada *Dataset*

Sebelum <i>Case Folding</i>	Setelah <i>Case Folding</i>
Sangat memudahkan masyarakat untuk akses dokumen kependudukan. Terima kasih Dirjen Dukcapil Kementerian Dalam Negeri	sangat memudahkan masyarakat untuk akses dokumen kependudukan terima kasih dirjen dukcapil kementerian dalam negeri
Aktivasinya sulit malah suruh ke dukcapil segala....	aktivasinya sulit malah suruh ke dukcapil segala

b. Tokenizing

Tokenizing adalah proses memecah teks menjadi unit-unit kecil yang atau token. Token bisa berupa kata, tanda baca, atau entitas penting lainnya. Pada tabel 4 berikut adalah contoh penggunaan *tokenizing* pada penelitian ini.

Tabel 4. Contoh *Tokenizing* pada *Dataset*

Sebelum <i>Tokenizing</i>	Setelah <i>Tokenizing</i>
sangat memudahkan masyarakat untuk akses dokumen kependudukan terima kasih dirjen dukcapil kementerian dalam negeri	['sangat', 'memudahkan', 'masyarakat', 'untuk', 'akses', 'kependudukan', 'terima', 'kasih', 'dirjen', 'dukcapil', 'kementerian', 'dalam', 'negeri']
aktivasinya sulit malah suruh ke dukcapil segala	['aktivasinya', 'sulit', 'malah', 'suruh', 'ke', 'dukcapil', 'segala']

c. Filtering (*Stopword Removal*)

Filtering adalah suatu proses untuk menghilangkan kata-kata umum dan tidak penting pada *dataset*. Kata-kata seperti "juga," atau," dan kata-kata serupa lainnya yang sering muncul dianggap sebagai *stopword*. Berikut pada tabel 5 merupakan contoh penerapan *Filtering* yang dilakukan pada penelitian ini.

Tabel 5. Contoh *Filtering* pada *Dataset*

Sebelum <i>Filtering</i>	Setelah <i>Filtering</i>
['sangat', 'memudahkan', 'masyarakat', 'untuk', 'akses', 'dokumen', 'kependudukan', 'terima', 'kasih', 'dirjen', 'dukcapil', 'kementrian', 'dalam', 'negeri']	['memudahkan', 'masyarakat', 'akses', 'dokumen', 'kependudukan', 'terima', 'kasih', 'dirjen', 'dukcapil', 'kementrian', 'negeri']
['aktivasinya', 'sulit', 'malah', 'suruh', 'ke', 'dukcapil', 'segala']	['aktivasinya', 'sulit', 'suruh', 'dukcapil']

d. *Stemming*

Stemming merupakan proses mengurangi kata-kata ke bentuk dasarnya atau akarnya. Tujuannya adalah untuk menyederhanakan berbagai bentuk kata menjadi satu bentuk dasar, misalnya mengubah "berlari," "berlari-lari," dan "pelari" menjadi "lari". Berikut pada tabel 6 merupakan contoh penerapan *Stemming* terhadap *dataset* yang dilakukan terhadap penelitian ini.

Tabel 6. Contoh *Stemming* pada *Dataset*

Sebelum <i>Stemming</i>	Setelah <i>Stemming</i>
['memudahkan', 'masyarakat', 'akses', 'dokumen', 'kependudukan', 'terima', 'kasih', 'dirjen', 'dukcapil', 'kementrian', 'negeri']	['mudah', 'rakyat', 'akses', 'dokumen', 'penduduk', 'terima', 'kasih', 'dirjen', 'dukcapil', 'mentri', 'negeri']
['aktivasinya', 'sulit', 'suruh', 'dukcapil']	['aktiv', 'sulit', 'suruh', 'dukcapil']

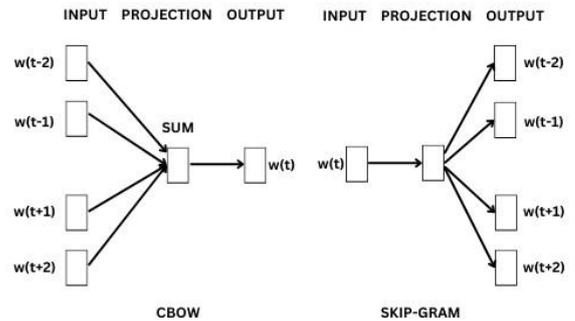
2.1.4 Pembagian Data

Data ulasan yang telah melewati tahap pra-proses data dan menjadi data bersih. kemudian dilakukan pembagian terhadap *sentiment* positif dan negatif dengan jumlah data yang seimbang. Ini perlu dilakukan agar tidak mengalami *imbalanced* pada data yang akan di gunakan. Metode yang digunakan dalam menyeimbangkan data adalah dengan menggunakan *Undersampling* terhadap data sentimen yang lebih besar, dalam hal ini data salah satu sentimen yang lebih besar akan dikurangkan menjadi seimbang keduanya. *Dataset* yang telah diproses tersisa 10800 data dengan pembagian data dengan *sentiment* positif 5400 dan dengan *sentiment* negatif 5400.

2.2 Pembentukan Model Word2Vec

Model Word2Vec adalah sebuah *model word embedding* yang diperkenalkan oleh Mikolov pada tahun 2013. Model ini mampu memahami makna semantik dari kata-kata dalam bahasa alami, sehingga dapat merepresentasikan setiap kata sebagai sebuah vektor. Pada penelitian ini, Word2Vec di bangun menggunakan *library* dari *Gensim* dengan menggunakan *corpus* yang di ambil dari *Wikipedia* sebanyak 200.000 artikel berbahasa Indonesia. Model Word2Vec dibangun dengan variasi CBOW dan Skip-Gram

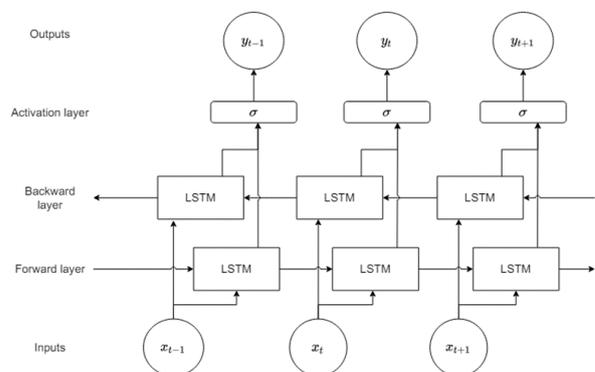
dengan jumlah dimensi vektor masing-masing 100, 200 dan 300. CBOW bekerja dengan memanfaatkan konteks untuk mengetahui target kata, memiliki kelebihan waktu *training* yang lebih cepat dan memiliki keuntungan untuk memprediksi kata-kata yang sering muncul. Sementara itu, Skip-Gram menggunakan satu kata untuk memprediksi konteks target dan memiliki kelebihan saat melakukan pemrosesan pada *dataset* yang relatif lebih sedikit, Skip-Gram mampu mengolah kata-kata yang jarang muncul dengan lebih baik. Pada gambar 4 di bawah ini merupakan arsitektur dari Word2Vec dengan 2 variasi CBOW dan Skip-Gram.



Gambar 4. Arsitektur Pembentukan Word2Vec CBOW dan Skip-Gram

2.3 Pembentukan *Bidirectional Long Short-Term Memory (Bi-LSTM)*

Bidirectional Long Short Term Memory (Bi-LSTM) adalah salah satu varian dari *Long Short Term Memory (LSTM)* yang telah dikembangkan lebih lanjut. Bi-LSTM menerima dua jenis masukan: *input forward* dan *input backward*. Hasil keluaran dari kedua *input* tersebut kemudian digabungkan. Dengan demikian, model mampu mempelajari informasi dari masa lalu dan masa depan untuk setiap urutan *input* yang diberikan. Kemampuan pemrosesan secara *bidirectional* memungkinkan model untuk memproses konteks tidak hanya konteks masa lalu tetapi juga masa depan, yang mengarah pada prediksi yang lebih terinformasi dan peningkatan akurasi dalam analisis sentimen[20]. Berikut pada gambar 5 di bawah ini merupakan Arsitektur dari *Bidirectional Long Short-Term Memory (Bi-LSTM)*.



Gambar 5. Arsitektur Bi-LSTM[21]

Bi-LSTM merupakan arsitektur LSTM yang terdiri dari dua lapisan LSTM yang berjalan paralel. Satu lapisan memproses sekuens dari awal ke akhir (*forward LSTM*), dan lapisan lainnya memproses sekuens dari akhir ke awal (*backward LSTM*). Hasil dari kedua lapisan kemudian digabungkan untuk menghasilkan representasi akhir.

Forward LSTM:

$$ht^{\rightarrow} = LSTM(xt, ht^{\rightarrow}1) \quad (1)$$

Backward LSTM :

$$ht^{\leftarrow} = LSTM(xt, ht^{\leftarrow} + 1) \quad (2)$$

Keluaran dari Bi-LSTM adalah kombinasi dari kedua arah:

$$ht = [ht^{\rightarrow}, ht^{\leftarrow}] \quad (3)$$

Di mana ht^{\rightarrow} adalah *hidden state* dari *forward LSTM* dan ht^{\leftarrow} adalah *hidden state* dari *backward LSTM*.

2.4 Skenario Penelitian

Penelitian ini mengevaluasi kinerja model analisis sentimen menggunakan algoritma Bi-LSTM dengan kombinasi metode *embedding* kata Word2Vec (CBOW dan Skip-Gram) dan proporsi pembagian data pelatihan dan pengujian. Penelitian ini akan melakukan sebanyak 18 kali pengujian terhadap kombinasi dari variabel berikut.

1. Algoritma Bi-LSTM
2. Word2Vec: CBOW dan Skip-Gram
3. Ukuran dimensi *word vector* CBOW : 100, 200, dan 300
4. Ukuran dimensi *word vector* Skip-Gram : 100, 200, dan 300
5. Proporsi pembagian data: 70/30, 80/20, dan 90/10

Selanjutnya, hasil dari penelitian ini akan dibandingkan dengan pengujian dengan beberapa metode dari *machine learning* klasik, seperti *Naïve Bayes*, *LGB Machine*, dan *Gradient Boosting*.

2.5 Evaluasi Model

Pada evaluasi model, Kinerja model akan diukur melalui serangkaian metrik, seperti *accuracy*, *recall*, *precision*, dan *F1-score*. Penggunaan evaluasi model ini sangat penting dalam menentukan seberapa baik model tersebut dalam melakukan prediksi berdasarkan data yang diberikan[22]. Data latih yang digunakan dalam proses pembelajaran model ini merupakan kumpulan data yang memungkinkan model mengenali pola dan menghasilkan prediksi. Data uji yang berbeda dari data latih digunakan untuk menguji kemampuan model yang telah dilatih. Data uji ini membantu dalam mengevaluasi seberapa baik model dapat menggeneralisasi dan memberikan prediksi yang akurat pada data yang belum pernah dilihat sebelumnya. *Confusion Matrix* (CM) akan digunakan dalam tahap evaluasi model untuk memberikan analisis yang lebih mendalam tentang performa model. *Confusion Matrix* (CM) ini berbentuk tabel

yang akan digunakan untuk menghitung dan mengevaluasi performa model yang diuji. Pada tabel 7 di bawah ini merupakan contoh dari penghitungan *Confusion Matrix* yang digunakan untuk melakukan *Evaluasi Model*.

Tabel 7. *Confusion Matrix* Untuk Evaluasi Model

		Actual Values	
		Positif (1)	Negatif (0)
Predicted Values	Positif (1)	TP (True Positif)	FP (False Positif)
	Negatif (0)	FN (False Negative)	TN (True Negative)

- a. *True Positif* (TP) adalah jumlah data yang memang positif dan diprediksi positif oleh model.
- b. *False Positif* (FP) adalah jumlah data yang sebenarnya negatif tetapi diprediksi positif oleh model.
- c. *False Negative* (FN) adalah Jumlah data yang sebenarnya positif tetapi diprediksi negatif oleh model.
- d. *True Negative* (TN) adalah Jumlah data yang memang negatif dan diprediksi negatif oleh model.

Berikut persamaan dari *confusion matrix* yang digunakan sebagai perhitungan performa.

$$Accuracy (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (4)$$

$$Recall (\%) = \frac{TP}{TP + FN} \times 100\% \quad (5)$$

$$Precision (\%) = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

$$F1-Score (\%) = \frac{2 \times precision \times recall}{precision + recall} \times 100\% \quad (7)$$

3. HASIL DAN PEMBAHASAN

Penelitian ini melakukan pengujian terhadap keseluruhan variabel dengan berbagai kombinasi menghasilkan 18 kali pengujian.

Hasil Pengujian Kombinasi Bi-LSTM dan Skip-Gram

Pada tabel 8, merupakan hasil dari pengujian menggunakan kombinasi antara Bi-LSTM dan Word2Vec dengan variasi Skip-Gram.

Tabel 8. Hasil Pengujian Dengan kombinasi Bi-LSTM dan Word2Vec Skip-Gram

Proporsi Pembagian Data	Dimensi Vektor	Akurasi
70/30	100	94,72 %
70/30	200	95,03 %
70/30	300	95,03 %
80/20	100	95,51 %
80/20	200	94,60 %

Proporsi Pembagian Data	Dimensi Vektor	Akurasi
80/20	300	95,65 %
90/10	100	95,00 %
90/10	200	94,81 %
90/10	300	94,63 %

Hasil Pengujian Kombinasi CBOW

Tabel 9 berikut merupakan hasil dari pengujian menggunakan kombinasi antara Bi-LSTM dan Word2Vec dengan variasi Word2Vec.

Tabel 9. Hasil Pengujian Dengan kombinasi Bi-LSTM dan Word2Vec Skip-Gram

Proporsi Pembagian Data	Dimensi Vektor	Akurasi
70/30	100	95,06 %
70/30	200	95,46 %
70/30	300	95,15 %
80/20	100	95,74 %
80/20	200	96,06 %
80/20	300	95,56 %
90/10	100	94,81 %
90/10	200	94,72 %
90/10	300	94,91 %

Word Embedding

Setelah penelitian dilaksanakan, kedua metode *embedding* kata, yaitu Skip-Gram dan CBOW menunjukkan performa yang baik dengan akurasi di atas 94% pada semua kombinasi. CBOW cenderung memberikan hasil yang sedikit lebih baik dalam beberapa kombinasi dibandingkan dengan Skip-Gram. Pada proporsi pembagian data 80/20 dengan dimensi vektor 200, CBOW mencapai akurasi

tertinggi sebesar 96,06%, sedangkan Skip-Gram mencapai akurasi tertinggi sebesar 95,65% pada proporsi dan dimensi yang sama.

Pengaruh Ukuran Dimensi Vektor

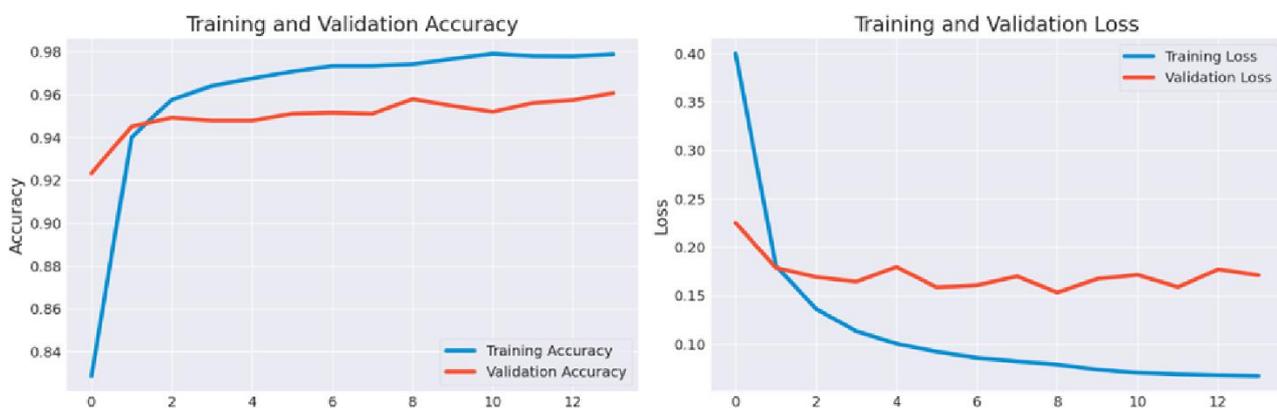
Pada metode Skip-Gram, akurasi sedikit bervariasi dengan perubahan dimensi vektor. Contohnya, pada proporsi pembagian data 70/30, akurasi meningkat dari 94,72% (dimensi 100) menjadi 95,03% (dimensi 200), namun tidak ada peningkatan lebih lanjut ketika dimensi dinaikkan menjadi 300. Pada metode CBOW juga tidak mengalami peningkatan akurasi yang konsisten dengan peningkatan dimensi vektor.

Pengaruh Proporsi Pembagian Data

Pada metode Skip-Gram, proporsi pembagian data 80/20 dengan dimensi vektor 300 memberikan akurasi tertinggi sebesar 95,65%. Kemudian pada metode CBOW, proporsi pembagian data 80/20 dengan dimensi vektor 200 juga memberikan akurasi tertinggi sebesar 96,06%. Untuk kedua metode, proporsi pembagian data 80/20 cenderung memberikan akurasi yang lebih tinggi dibandingkan dengan proporsi 70/30 dan 90/10, menunjukkan bahwa proporsi data pelatihan yang optimal berkontribusi terhadap kinerja model yang lebih baik.

Setelah dilaksanakan penelitian, didapati model terbaik dari *Bidirectional Long Short Term Memory* dan Word2Vec untuk pengembangan analisis sentimen aplikasi Identitas Kependudukan Digital. Kombinasi Word2Vec CBOW dengan pembagian data 80/20 dan ukuran vektor sebesar 200 dimensi menjadi model terbaik pada penelitian ini.

Berikut pada gambar 6 merupakan grafik dari hasil *Training Accuracy* dan *Validation Accuracy* model CBOW dengan ukuran *Word Vector* 200 dan pembagian data 80/20.



Gambar 6. Grafik *Training and Validation Accuracy* Model CBOW 200 dengan pembagian data 80/20

Pada tabel 10 merupakan hasil dari nilai evaluasi *Confusion Matrix* pada Model terbaik dengan kombinasi Variasi CBOW dengan ukuran dimensi Word Vector 200 dan Pembagian Data 80/20.

Tabel 10. Hasil *Confusion Matrix*

Predicted Values	Actual Values	Actual Values	
		Positif (1)	Negatif (0)
Positif (1)	Positif (1)	1032	47
	Negatif (0)	38	1043

Pada tabel 11 di bawah ini merupakan hasil dari *Confusion Matrix* yang telah di hitung, sehingga memperoleh hasil *Accuracy*, *Precision*, *Recall*, dan *F1-Score* terhadap Model CBOW 200 dan Pembagian Data 80/20.

Tabel 11. Hasil Evaluasi dari *Confusion Matrix*

Confusion Matrix Evaluation			
Accuracy	Precision	Recall	F1-Score
96,06	96,44	95,64	96,04

Hasil dari Pengujian terhadap Metode *Machine Learning* Klasik

Penelitian ini juga membandingkan hasil yang diperoleh dari penerapan *Deep Learning* melalui Bi-LSTM dan Word2Vec dengan penerapan *Machine learning* klasik. Dalam penelitian ini menggunakan beberapa algoritma seperti Naive Bayes, *LGB Machine*, dan *Gradient Boosting* sebagai pembanding dari hasil penelitian. Berikut pada tabel 12 merupakan hasil dari penggunaan *Machine Learning* klasik terhadap ulasan pengguna aplikasi Identitas Kependudukan Digital.

Tabel 12. Hasil Pengujian dengan Metode *Machine Learning* Klasik

Model	Accuracy	Roc Scores
Naive Bayes	93,07	93,05
LGB Machine	90,70	91,71
Gradient Boosting	86,33	86,25

Hasil dari pengujian terhadap metode *machine learning*, model Naive Bayes mendapatkan nilai akurasi tertinggi dengan 93,07%, melebihi kedua algoritma *machine learning* yang lain. Namun hasil tertinggi dari metode *machine learning* klasik ini belum bisa melebihi hasil dari kombinasi Bi-LSTM dan Word2Vec keseluruhan.

Secara keseluruhan, penelitian ini mengevaluasi kinerja model analisis sentimen menggunakan algoritma Bi-LSTM dan Word2Vec untuk mengklasifikasikan opini publik terhadap aplikasi Identitas Kependudukan Digital (IKD).

Evaluasi dilakukan berdasarkan nilai-nilai dalam *confusion matrix*, yang menunjukkan performa model dalam membedakan antara sentimen positif dan negatif. Hasil menunjukkan bahwa model terbaik mencapai *accuracy* sebesar 96,06%, *precision* 96,44%, *recall* 95,64%, dan *f1-score* 96,04%. Dengan akurasi model yang tinggi menunjukkan kinerja yang baik dalam mengklasifikasikan sentimen dalam teks terkait aplikasi IKD. Pengembangan analisis sentimen ini dapat membantu berbagai pihak seperti Pemerintah dan pengembang dalam memahami opini publik. Dengan memahami sentimen publik, pemerintah dan pengembang aplikasi IKD dapat membuat keputusan yang lebih tepat sasaran dan bermanfaat bagi masyarakat luas dengan memberikan pengalaman pengguna yang lebih baik. Hasil penelitian ini cukup baik jika dibandingkan dengan penelitian sebelumnya. Misalnya, penelitian oleh [2] yang menggunakan algoritma *Naive Bayes* untuk analisis sentimen terhadap aplikasi IKD, dan penelitian oleh [13] yang menggunakan algoritma LSTM dan Word2Vec untuk analisis sentimen terhadap ulasan film. Perbandingan ini menunjukkan keunggulan penggunaan Bi-LSTM dan Word2Vec dalam analisis sentimen, khususnya untuk aplikasi IKD. Secara keseluruhan, penelitian ini memberikan kontribusi dalam bidang analisis sentimen dengan menggunakan pendekatan yang berbeda dari penelitian sebelumnya.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, penerapan metode *Bidirectional LSTM* dengan Word2Vec dapat melakukan analisis sentimen terhadap data ulasan aplikasi Identitas Kependudukan Digital dengan sangat baik dan melebihi akurasi dari penelitian terkait aplikasi Identitas Kependudukan Digital sebelumnya yang menggunakan metode *Machine Learning* klasik. Hasil dari penelitian ini memenuhi kebutuhan akan penerapan metode *Deep Learning* terhadap ulasan aplikasi Identitas Kependudukan Digital. Pada pengujian terhadap ulasan aplikasi Identitas Kependudukan Digital model kombinasi dari Bi-LSTM dan Word2Vec variasi CBOW mendapatkan akurasi rata-rata lebih tinggi dibandingkan dengan kombinasi dari Bi-LSTM dan Word2Vec variasi Skip-Gram. Kemudian pembagian data dengan proporsi 80/20 mendapatkan hasil rata-rata lebih tinggi dibandingkan dengan pembagian data 70/30 dan 90/10. Implikasi dari penelitian ini dijadikan sebagai pengembangan analisis sentimen dari penelitian sebelumnya yang menggunakan metode *machine learning* terhadap aplikasi Identitas Kependudukan Digital dan sebagai perbandingan dengan penelitian-penelitian sebelumnya. Untuk penelitian selanjutnya, disarankan untuk membandingkan Word2Vec dengan metode *word embedding* lainnya seperti *FastText*, serta menggunakan data yang lebih besar dan relevan dengan topik analisis sentimen yang akan dikerjakan, guna meningkatkan akurasi dan relevansi hasil penelitian.

DAFTAR PUSTAKA

- [1] A. Yulanda and M. Fachri Adnan, "Transformasi Digital: Meningkatkan Efisiensi Pelayanan Publik Ditinjau dari Perspektif Administrasi Publik," 2023. [Online]. Available: <https://isora.tpublising.org/index.php/isora>
- [2] R. Hidayat, R. Nur Rahman, M. Reifin Perdana, P. Teknik Informatika, F. Sains dan Teknologi, and U. Muhammadiyah Kalimantan Timur, "Analisis Sentimen Aplikasi Identitas Kependudukan Digital (IKD) Menggunakan Metode Naïve Bayes," *Jurnal Sistem Informasi dan Ilmu Komputer*, vol. 2, no. 1, pp. 129–140, 2024, doi: 10.59581/jusiik-widyakarya.v2i1.2320.
- [3] L. M. Azizah, D. B. Ajipratama, N. A. R. Putri, and C. Damarjati, "Analisa Sentimen Masyarakat Terhadap Kebijakan Vaksinasi Covid-19 Di Indonesia Pada Twitter Menggunakan Algoritma LSTM La," *JURNAL IPTEKKOM Jurnal Ilmu Pengetahuan & Teknologi Informasi*, vol. 24, no. 2, pp. 161–172, Dec. 2022, doi: 10.17933/iptekkom.24.2.2022.161-172.
- [4] R. Maulana, A. Voutama, and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC," *Jurnal Teknologi Terpadu*, vol. 9, no. 1, 2023, doi: 10.54914/jtt.v9i1.609.
- [5] P. SV *et al.*, "Twitter-Based Sentiment Analysis and Topic Modeling of Social Media Posts Using Natural Language Processing, to Understand People's Perspectives Regarding COVID-19 Booster Vaccine Shots in India: Crucial to Expanding Vaccination Coverage," *Vaccines (Basel)*, vol. 10, no. 11, Nov. 2022, doi: 10.3390/vaccines10111929.
- [6] A. Komarudin and A. M. Hilda, "Analisis Sentimen Ulasan Aplikasi Identitas Kependudukan Digital Pada Play Store Menggunakan Metode Naïve Bayes," *Computer Science (CO-SCIENCE)*, vol. 4, no. 1, 2024, doi: 10.31294/coscience.v4i1.2955.
- [7] R. A. Lestari, A. Erfina, and W. Jatmiko, "Penerapan Algoritma Support Vector Machine pada Analisis Sentimen Terhadap Identitas Kependudukan Digital," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 5, 2023, doi: 10.25126/jtiik.20231057264.
- [8] V. Umarani, A. Julian, and J. Deepa, "Sentiment Analysis using various Machine Learning and Deep Learning Techniques," *Journal of the Nigerian Society of Physical Sciences*, vol. 3, no. 4, 2021, doi: 10.46481/jnsps.2021.308.
- [9] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 8, no. 4, 2018, doi: 10.1002/widm.1253.
- [10] A. Rahman *et al.*, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen".
- [11] M. Z. Rahman, Y. A. Sari, and N. Yudistira, "Analisis Sentimen Tweet COVID-19 menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 11, pp. 5120–5127, 2021, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [12] A. F. Hidayatullah, S. Cahyaningtyas, and A. M. Hakim, "Sentiment Analysis on Twitter using Neural Network: Indonesian Presidential Election 2019 Dataset," *IOP Conf Ser Mater Sci Eng*, vol. 1077, no. 1, 2021, doi: 10.1088/1757-899x/1077/1/012001.
- [13] W. Widayat, "Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 3, p. 1018, Jul. 2021, doi: 10.30865/mib.v5i3.3111.
- [14] P. F. Muhammad, R. Kusumaningrum, and A. Wibowo, "Sentiment Analysis Using Word2vec and Long Short-Term Memory (LSTM) for Indonesian Hotel Reviews," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 728–735. doi: 10.1016/j.procs.2021.01.061.
- [15] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment texts based on BiLSTM," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2909919.
- [16] D. I. Af'idah, D. Dairoh, S. F. Handayani, R. W. Pratiwi, and S. I. Sari, "Sentimen Ulasan Destinasi Wisata Pulau Bali Menggunakan Bidirectional Long Short Term Memory," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 3, pp. 607–618, Jul. 2022, doi: 10.30812/matrik.v21i3.1402.
- [17] J. Zhou, Y. Lu, H. N. Dai, H. Wang, and H. Xiao, "Sentiment analysis of Chinese microblog based on stacked bidirectional LSTM," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2905048.
- [18] F. A. Nugraha, N. H. Harani, R. Habibi, and Rd. N. S. Fatonah, "Sentiment Analysis on Social Distancing and Physical Distancing on Twitter Social Media using Recurrent Neural Network

- (RNN) Algorithm,” *Jurnal Online Informatika*, vol. 5, no. 2, p. 195, Dec. 2020, doi: 10.15575/join.v5i2.632.
- [19] D. F. Sengkey, A. Jacobus, and F. J. Manoppo, “Effects of kernels and the proportion of training data on the accuracy of svm sentiment analysis in lecturer evaluation,” *IAES International Journal of Artificial Intelligence*, vol. 9, no. 4, 2020, doi: 10.11591/ijai.v9.i4.pp734-743.
- [20] D. Shrimal, L. Suganya, and P. Deshmukh, “*Kanak Pandit 2 Harshali Patil Regular paper Comparative Analysis of Deep Learning Models for Sentiment Analysis on IMDB Reviews,” 2024.
- [21] S. Mikhailov and A. Kashevnik, “Car tourist trajectory prediction based on bidirectional lstm neural network,” *Electronics (Switzerland)*, vol. 10, no. 12, 2021, doi: 10.3390/electronics10121390.
- [22] O. Majdoubi, A. Benba, and A. Hammouch, “Comprehensive Machine Learning And Deep Learning Approaches For Parkinson’s Disease Classification And Severity Assessment,” *Informatyka, Automatyka, Pomiar w Gospodarce i Ochronie Srodowiska*, vol. 13, no. 4, 2023, doi: 10.35784/iapgos.5309.