

Analisis Sentimen Masyarakat Terhadap Konten Bahasa Isyarat Pada Media Sosial Tiktok Menggunakan RNN dan BERT

Christian Sri Kusuma Aditya*¹, Asfa Maghfiratunnisa²

^{1,2}Universitas Muhammadiyah Malang

christianskaditya@umm.ac.id¹, asfamaghfiratunnisa@gmail.com²

Abstrak

Penelitian ini bertujuan untuk mengeksplorasi sentimen masyarakat terhadap konten bahasa isyarat di media sosial Tiktok menggunakan pendekatan deep learning dengan model RNN(LSTM) dan BERT. Data dikumpulkan, diproses, dan dilabeli untuk membentuk dataset yang digunakan dalam analisis. Hasil klasifikasi sentimen dari model RNN(LSTM) menunjukkan tingkat akurasi sebesar 0.90%, menunjukkan efektivitas model dalam membedakan sentimen antara dua kelas. Sementara itu, model BERT mencapai tingkat akurasi sebesar 0.80%, menunjukkan kemampuan model dalam mengklasifikasikan sentimen dari data yang diberikan. Penggabungan model RNN dan BERT menunjukkan peningkatan kinerja dalam mengklasifikasikan sentimen, bahwa pendekatan gabungan dapat meningkatkan pemahaman model terhadap konten yang kompleks. Selain itu, kombinasi model BERT dengan RNN juga terbukti efektif dalam mengatasi kompleksitas analisis sentimen dalam konteks bahasa isyarat di media sosial. Evaluasi berkelanjutan dan validitas model yang melibatkan ahli dalam Bahasa isyarat direkomendasikan untuk memastikan keberlanjutan dan validitas model yang dikembangkan. Penelitian ini diharapkan dapat memberikan kontribusi penting dalam memahami pengaruh media sosial terhadap pandangan masyarakat terhadap bahasa isyarat dan disabilitas pendengaran secara umum, serta memperkuat upaya menuju masyarakat yang lebih inklusif.

Kata Kunci: Bahasa isyarat, Analisis sentimen, Tiktok, RNN(LSTM), BERT, Deep learning

Abstract

This research aims to explore public sentiment towards sign language content on Tiktok social media using a deep learning approach with RNN(LSTM) and BERT models. Data is collected, processed, and labeled to form a dataset used in analysis. The sentiment classification results from the RNN(LSTM) model show an accuracy level of 0.90%, indicating the effectiveness of the model in distinguishing sentiment between the two classes. Meanwhile, the BERT model achieved an accuracy level of 0.80%, showing the model's ability to classify sentiment from the data provided. Combining RNN and BERT models shows improved performance in classifying sentiment, that the combined approach can improve the model's understanding of complex content. In addition, the combination of the BERT model with RNN has also proven effective in overcoming the complexity of sentiment analysis in the context of sign language on social media. Continuous evaluation and validity of the model involving experts in sign language is recommended to ensure the sustainability and validity of the developed model. It is hoped that this research will make an important contribution to understanding the influence of social media on society's views of sign language and hearing disabilities in general, as well as strengthening efforts towards a more inclusive society.

Keywords: Sign Language, Sentiment Analysis, TikTok, RNN(LSTM), BERT, Deep Learning

1. Pendahuluan

Media sosial merupakan teknologi komputer yang dirancang untuk memudahkan komunikasi dengan orang lain tanpa harus bertemu langsung, untuk bersenang-senang dan mengurangi perasaan terisolasi [1]. Salah satu aplikasi media sosial yang menjadi populer sejak tahun 2020 adalah Tiktok. Belakangan ini, aplikasi Tiktok sedang menjadi trending di kalangan masyarakat umum. Pada aplikasi Tiktok masyarakat dapat menikmati konten berupa musik, dance, video kreasi ataupun cuplikan drama atau film. Drama Korea atau K-Drama mempunyai basis penggemar yang besar di seluruh dunia, termasuk di dalamnya yaitu kalangan netizen

Tiktok. Pengguna berbagai konten terkait drama, seperti lip-syncing, reenactments, atau komentar, sehingga hal ini dapat menciptakan interaksi yang intens di antara mereka. Salah satu pengguna yang ikut terlibat di dalamnya adalah orang penyandang disabilitas tuli atau bisu. Bahasa isyarat mungkin digunakan dalam cerita untuk menyampaikan pesan atau menggambarkan karakter dengan kebutuhan khusus.

Menurut Organisasi Kesehatan Dunia, 5% dari populasi dunia sekitar 466 juta orang tuli atau bisu atau memiliki gangguan pendengaran yang melumpuhkan [2]. Seringkali ada tembok perbedaan antara penyandang disabilitas dan non-disabilitas dalam menikmati sosial media Tiktok. Sentimen analisis dapat membantu memahami bagaimana masyarakat pengguna aplikasi tiktok merespon elemen-elemen tertentu dalam drama korea, terutama jika bahasa isyarat menjadi fokus penceritaan. Pengguna Tiktok sering menggunakan hastag untuk menyatukan konten terkait. Penelitian ini dapat membantu memahami apakah respons netizen di TikTok terhadap adegan bahasa isyarat dalam drama Korea bersifat positif, negatif, atau netral. Dengan mendokumentasikan respons dan diskusi netizen, analisis sentimen dapat membantu mengidentifikasi sejauh mana representasi ini diterima dan diapresiasi oleh komunitas di media Tiktok dan mengatasi stigma yang mungkin terkait dengan komunitas penyandang disabilitas.

Kombinasi antara metode deep learning RNN LSTM dan pendekatan dari pre-trained BERT merupakan strategi yang inovatif dalam melakukan analisis teks yang kompleks. RNN LSTM yang merupakan sebuah teknik dalam deep learning memiliki kelebihan dalam memproses data sekuensial termasuk data teks, dengan kemampuan untuk menangkap hubungan temporal antar kata serta memertahankan informasi jangka Panjang [3]. Namun, model RNN LSTM sering menghadapi kendala seperti vanishing atau exploding gradient, serta kesulitan dalam menangani konteks teks yang sangat kompleks [4]. Di sisi lain, BERT (Bidirectional Encoder Representations from Transformers) adalah sebuah model bahasa yang dilatih secara unsupervised untuk memahami bahasa alami [5]. Dengan menggabungkan kedua metode ini, RNN LSTM dapat memanfaatkan kemampuan BERT dalam memahami konteks kata secara menyeluruh, sementara BERT dapat mengatasi kendala-kendala yang sering dihadapi oleh RNN LSTM.

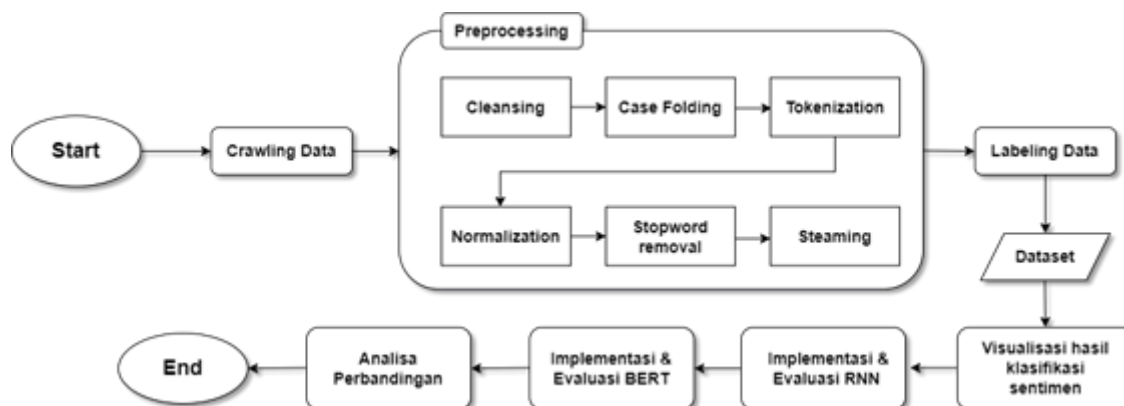
Penelitian yang berfokus pada analisis sentimen di platform Tiktok, menyelusuri berbagai metode dan pendekatan untuk memahami pola sentimen pengguna media sosial. Seperti yang digunakan oleh Siswanto, et al (2022) terkait penggunaan fitur berbasis lexicon sebagai pelabelan data review Tiktok yang dapat mengklasifikasi sentimen menjadi positif dan negatif tanpa pelabelan manual sehingga dapat menghemat waktu dan memberikan performa yang lebih baik dibandingkan dengan tanpa fitur berbasis lexicon [6]. Begitu juga dengan Riky

Iskandar Syah, et al (2023) yang juga menggabungkan metode Lexicon Based dan Naïve Bayes dimana lexicon digunakan untuk pelabelan dan Naïve Bayes digunakan untuk klasifikasi [7]. Dari hasil tersebut pengkombinasian metode Lexicon Based dan Naïve Bayes ini memiliki akurasi yang sangat tinggi pada media sosial Tiktok dan Youtube. Selain itu, penelitian yang dilakukan oleh R.A.Pramunendar, et al (2022) mengusulkan model RNN-LSTM yang menerapkan arsitektur Bidirectional LSTM layer dengan variational dropout [4]. Berdasarkan penelitian ini dapat disimpulkan bahwa model terbaik untuk analisis sentimen opini terkait Vaksin COVID-19 dalam penelitian ini adalah model Variational Bi-LSTM. Adapun penelitian lain yang menggunakan platform media sosial Tiktok dilakukan oleh Firman Nurdiansyah, et al (2023)[8]. Berdasarkan penelitian analisis sentimen terhadap opini publik pada platform Tiktok menggunakan metode Long Short-Term Memory (LSTM), dengan sebuah dataset yang terdiri dari 1550 data, telah diperoleh hasil akurasi yang mencapai tingkat yang optimal yaitu 100%. Penelitian selanjutnya oleh Tosti H. C. Cshiang, dkk (2022) dengan penelitian yang berjudul Investigating the Difference of Fake News Source Credibility Recognition between ANN and BERT Algorithms in Artificial Intelligence, dari penelitian tersebut peneliti berhasil mendapatkan hasil akurasi ANN 82.75% dan BERT 91.2% [9].

Namun, penelitian sebelumnya masih jarang membahas mengenai sentimen masyarakat terhadap konten bahasa isyarat pada media sosial Tiktok. Oleh karena itu peneliti akan berfokus melakukan pembaharuan yaitu mengenai konten bahasa isyarat pada media sosial Tiktok sebagai platform utama. Untuk dapat memberikan kontribusi yang signifikan pada pemahaman kita tentang bagaimana media sosial dapat memengaruhi pandangan masyarakat terhadap bahasa isyarat dan disabilitas pendengaran secara umum. Hal ini juga dapat memperkuat upaya menuju masyarakat yang lebih inklusif dan sadar akan keberagaman.

2. Metode Penelitian

2.1 Rancangan Penelitian



Gambar 1. Rancangan Penelitian

2.2 Data Crawling

Data untuk penelitian ini diperoleh dari hasil Crawling 12.915 respon netizen Tiktok mengenai adegan bahasa isyarat dalam drama korea *Twinkling Watermelon* dengan hastag #twinklingwatermelon, #handsignlanguage, #signlanguage dan #bahasaisyarat sejak tanggal 1 Desember hingga 10 Desember 2023 dari beberapa konten di aplikasi Tiktok. Proses crawling data dilakukan menggunakan Python dan scraper library Tiktok dengan layanan API yang dirancang khusus untuk mengambil data dari tiktok.

Tabel 1. Sample Dataset

No	Full_text	Label
1	is this malaysia	Non-influential
2	i want to learn sign butdont know any good source do you have any source for me	Influential
3	hyjiyn and they say kdrama isbad influence	Non-influential
4	i hope you will be givensmooth run	Influential
5	you are not deaf you are so special	Influential
6	really fypnow learn the sign language all come	Non-influential
7	beyond grateful to the person who founded sign language	Influential
8	qualnome onde posso assistant	Non-influential
9	shes so prettyyrit	Non-influential
10	thank you for teaching me the sign language now if the un is not afraid anymore	Influential

2.3 Preprocessing Data

Preprocessing data adalah serangkaian langkah-langkah yang dilakukan untuk membersihkan, penyortiran, konversi data mentah menjadi menjadi representasi logis untuk pemodelan dan analisis. Tujuan dari preprocessing data adalah untuk meningkatkan kualitas data, memperbaiki masalah yang mungkin akan terjadi, dan membuat data agar sesuai dengan penggunaan dalam algoritma atau model yang akan diterapkan.

2.3.1 Data Cleaning

Data cleaning merupakan langkah penting dalam preprocessing data yang bertujuan untuk membersihkan dan mempersiapkan data mentah untuk digunakan secara efektif dalam analisis dan pemodelan. Proses ini mencakup identifikasi dan penanganan masalah yang mungkin muncul dalam dataset seperti missing value, outliers, duplikasi dan lain-lain. Hal ini dapat membantu menghilangkan noise yang tidak diinginkan atau simbol yang tidak relevan dari teks.

2.3.2 Case Folding

Case folding merupakan suatu proses mengonversi semua karakter dalam sebuah teks ke bentuk yang seragam, biasanya dengan mengubah seluruh teks menjadi huruf kecil atau huruf besar. Tujuan dari case folding ialah untuk menghilangkan variasi dalam penulisan karakter, sehingga memudahkan perbandingan dan analisis teks tanpa memperhatikan perbedaan huruf besar atau kecil.

2.3.3 Tokenization

Tokenisasi merupakan proses membagi sebuah teks atau kalimat menjadi unit-unit yang lebih kecil atau disebut dengan token. Token dapat berupa kata, frasa, atau katrakter tergantung pada tingkat granularitas yang diinginkan. Tokenisasi merupakan langkah penting dalam pemrosesan bahasa alami NLP (Natural language Processing) dan membantu dalam pemahaman struktur teks dan pengolahan informasi lebih lanjut. Tokenisasi ini diproses dengan menggunakan library dari NLTK yaitu `nlk.tokenize`.

2.3.4 Normalization

Normalisasi dalam konteks pemrosesan data merujuk pada serangkaian langkah untuk mengubah atau menyusun data sehingga memiliki format atau skala yang beragam. Tujuan dari normalisasi adalah mengubah kata-kata yang disingkat, tidak standar, atau salah eja menjadi kata standar. Normalisasi umumnya digunakan dalam berbagai bidang seperti statistika, machine learning dan data mining. Tujuan normalisasi adalah untuk mengubah data dari dataset menjadi bahasa standar dan formal.

2.3.5 Stopword Removal

Stopword removal merupakan penghapusan stopwords, termasuk kata-kata umum yang tidak membawa makna atau fungsi signifikan dalam konteks analisis. Contoh stopwords termasuk artikel, konjungsi dan reposisi. Selain itu, kamus stopwords juga ditambahkan berdasarkan kata-kata yang tidak diperlukan dalam dataset.

2.3.6 Stemming

Stemming merupakan proses mengurangi kata ke bentuk dasar atau akar mereka. Ini bertujuan untuk menghilangkan variasi infleksi dan membawa kata-kata terkait ke bentuk dasar umum. Ini membantu mengurangi redundansi dan mengkonsolidasikan kata-kata serupa.

2.4 Tahapan Scrum

Dari hasil crawling data respon masyarakat terhadap konten bahasa isyarat pada media sosial tiktok yang telah di dapat, tahap selanjutnya adalah membuat label menggunakan Lexicon. Sering kali sentimen sebuah komentar dari media sosial sulit untuk ditentukan secara benar ketika komentar tersebut mengandung unsur sarkasme [10]. Sarkasme merupakan bentuk khusus dari ironi yang terjadi ketika seseorang mengungkapkan sesuatu yang sebenarnya bertentangan dengan makna harfiahnya sehingga hal ini menjadi masalah yang sulit dalam proses analisis sentimen bahkan bagi manusia. Lexicon VADER dapat mengatasi komentar yang mengandung sarkasme dalam kalimatnya tanpa harus melalui pelabelan manual dari manusia karena memanfaatkan daftar kata-kata yang telah diakurasi dengan cermat [11]. Lexicon VADER telah dikembangkan dengan menggunakan lexicon sentimen yang dianggap sebagai standar emas untuk analisis sentimen dengan fitur dan karakteristik unik di media sosial yang seringkali menghasilkan teks yang pendek, informal dan kaya akan ekspresi sentimen[12]. Dengan menerapkan metode lexicon, teks dapat diidentifikasi dalam kategori positif, negatif atau netral. Namun, mengingat sifat data yang tidak terstruktur di media sosial, diperlukan fleksibilitas metodologis dan adaptasi [13]. Penggunaan analisis teks otomatis dapat memberikan gambaran komprehensif tentang dataset yang luas dan menghindari seleksi data kualitatif [14].

VADER (Valence Aware Dictionary and sEntiment Reasoner) merupakan sebuah algoritma analisis sentimen yang memanfaatkan pembobotan nilai kata-kata untuk menentukan nilai polaritas teks, baik itu positif, negatif atau netral. Jika nilai polaritas >1 maka opini dianggap positif, sebaliknya jika nilai polaritasnya <0 opini dianggap negatif, sedangkan jika nilai polaritas $= 0$, maka opini dianggap netral. Dalam menentukan nilai polaritas suatu teks, digunakan rumus [15] [Pamungkas & Putri, 2016 dikutip dalam Hamka & Sari, 2022].

$$Sentiments_{score} = \sum_{i=1}^n Sentiments_{score} + W_{positive} + W_{negative} \quad (1)$$

Analisis disusun berdasarkan respon pengguna terhadap konten bahasa isyarat, dimana label Influential diidentifikasi menggunakan sentimen positif dan negatif sedangkan label Non-influential diidentifikasi menggunakan sentimen netral dari pengguna Tiktok pada konten tersebut [16]. Penentuan polaritas sentimen pada suatu ulasan dilakukan dengan memanfaatkan persamaan [17][Vu & Le, 2017].

$$Sentiments_{score} \begin{cases} \text{positif,} & \text{jika } Sentiments_{score} > 0 \\ \text{netral,} & \text{jika } Sentiments_{score} = 0 \\ \text{negatif,} & \text{jika } Sentiments_{score} < 0 \end{cases} \quad (2)$$

Dalam penelitian ini, Valence Aware Dictionary and sEntiment Reasoner (VADER) dipilih sebagai alat analisis sentimen. VADER, sebuah leksikon sentimen sumber terbuka, mampu memahami elemen seperti emotikon dan bahasa gaul yang sering digunakan di media sosial. Lexicon Based merupakan sebuah pendekatan yang mencakup frasa, ekspresi, atau konten dalam bentuk teks yang biasa ditemukan di ruang obrolan, dialog, pesan, ulasan dan lainnya [7].

2.5 Visualisasi Klasifikasi

Visualisasi klasifikasi sentimen melibatkan representasi grafis dari hasil analisis sentimen pada teks. Ini bisa dilakukan dengan berbagai cara, salah satunya adalah menggunakan visualisasi seperti grafik batang, pie chart, atau word cloud untuk menunjukkan distribusi sentimen Influential dan Non-influential dalam dataset atau hasil klasifikasi. Visualisasi kemudian digunakan untuk mewakili hasil analisis ini dalam bentuk yang lebih mudah dipahami.

2.6 Implementasi dan Evaluasi RNN

Implementasi RNN (Recurrent Neural Network) melibatkan langkah-langkah yang penting dalam membangun model untuk menganalisis data sekuensial. RNN memiliki kemampuan untuk memodelkan ketergantungan antara kata-kata atau token dalam urutan khususnya pada respon pengguna Tiktok, dimana interaksi pengguna seringkali melibatkan komentar atau respon berturut-turut yang membentuk narasi atau konversasi. Salah satu metode deep learning yang diterapkan dalam penelitian ini adalah RNN dengan arsitektur Long Short-Term Memory (LSTM)[23]. Penggunaan model RNN dengan arsitektur LSTM untuk analisis sentimen merupakan pendekatan yang kuat dalam memahami dan mengklasifikasi sentimen dari teks dengan mempertimbangkan konteks temporal dari urutan tersebut. LSTM menggunakan unit memori sel untuk mempertahankan informasi jangka panjang dan mengatasi masalah vanishing gradient yang umum terjadi dalam RNN.

Penggunaan metrik evaluasi seperti akurasi, presisi, recall atau F1-score memberikan pemahaman yang mendalam tentang sejauh mana model mampu mengklasifikasikan atau memprediksi data yang belum pernah dilihat sebelumnya. Analisis kinerja melibatkan pemeriksaan hasil prediksi dan memahami kemampuan serta batasan model dalam menangani data sekuensial. RNN disebut memiliki memori yang dapat diingat karena hasilnya bergantung pada komputasi sebelumnya. Ini adalah perbedaan utama antara RNN dan jaringan saraf biasa, di mana RNN mampu menyimpan informasi dari komputasi sebelumnya dan menggunakannya pada elemen berikutnya dalam urutan masukan [6].

2.7 Implementasi dan Evaluasi BERT

Implementasi model BERT (Bidirectional Encoder Representations from Transformers) dalam analisis sentimen respon pengguna media sosial Tiktok melibatkan sejumlah langkah kunci untuk memahami dan mengartikan nuansa yang kompleks dalam teks. Proses fine-tuning memungkinkan model untuk menyesuaikan diri dengan bahasa dan gaya unik yang umumnya digunakan oleh pengguna Tiktok. Evaluasi BERT melibatkan beberapa metrik tergantung pada tugas yang dikerjakan. Misalnya, untuk tugas klasifikasi teks, metrik seperti akurasi, presisi, recall, dan F1-score sering digunakan.

Proses analisis sentimen dengan BERT dimulai dengan pemberian teks ke model. Selama pre-training, model dilatih dengan berbagai tugas pre-trained pada data yang tidak berlabel. Dilanjutkan dengan proses fine-tuning, yaitu proses inialisasi BERT dengan parameter pre-trained. Parameter sebelumnya akan dilatih ulang (fine-tuned) menggunakan data yang diberi label dari tugas turunan (downstream task)[18].

2.8 Evaluasi dan Perbaikan

Setelah menganalisis data, Langkah terakhir dalam metodologi penelitian ini adalah evaluasi game dan pelaksanaan perbaikan jika diperlukan. Evaluasi game melibatkan partisipasi dari tim pengembang, anak-anak, dan pendidik. Hasil evaluasi tersebut akan digunakan untuk memperbaiki game guna meningkatkan kualitas pengalaman belajar dan pemahaman keislaman yang dihasilkan. Seluruh tahapan ini membentuk kerangka kerja yang kokoh dan terperinci untuk pengembangan game edukasi yang efektif dan berorientasi pada kebutuhan anak-anak dalam pemahaman keislaman

3. Hasil dan Pembahasan

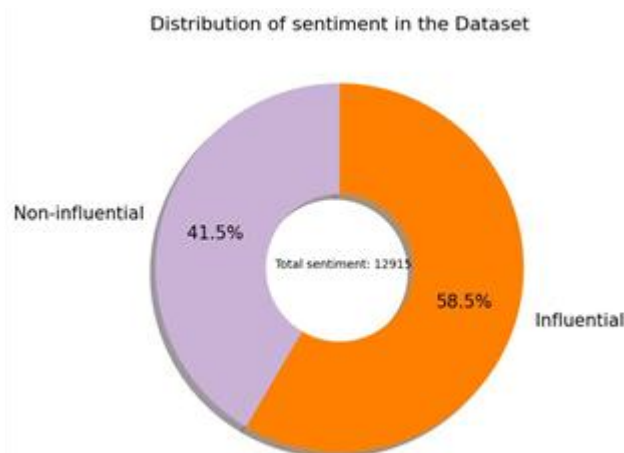
3.1 Pengujian Sistem

Proses pengujian sistem sedang dilaksanakan dengan melakukan evaluasi terhadap fungsionalitas dan kegunaannya. Pengujian fungsional lebih berfokus pada verifikasi apakah fitur-fitur beroperasi dengan baik sesuai fungsinya. Pengujian fungsional ini menggunakan metode black box testing, di mana pengujian dilakukan tanpa memerlukan pengetahuan terperinci tentang implementasi internal sistem. Melalui pengujian fungsional ini, dapat diidentifikasi kesalahan dalam fungsi, struktur, dan tampilan sistem.

Tabel 2. Data Preprocessing

No	User	Text Clean	Text Preprocessed
1	acakk	the influence of korean drama twinkling water melon on the desire to learn sign language	[influence, korean, drama, twinkling, water, melon, desire, learn, sign, language]
2	gisellalcantara0	this is the song in sign language that i do	[song,sign,language]
3	aftietaftie	i need to watch this drama	[need,watch,drama]

Tahap preprocessing, dilakukan penentuan polaritas untuk mengidentifikasi kata-kata yang mengandung sentimen negatif, positif, dan netral. Dalam penelitian ini, digunakan metode berbasis aturan untuk mengenali polaritas komentar dengan memanfaatkan kumpulan istilah yang sudah dikelompokkan berdasarkan polaritasnya (positif, negatif, atau netral). Analisis selanjutnya disusun berdasarkan tanggapan pengguna terhadap konten bahasa isyarat, dengan label "Influential" digunakan untuk mengidentifikasi sentimen positif dan negatif, sedangkan label "Non-influential" mengacu pada sentimen netral dari pengguna Tiktok terhadap konten tersebut.



Gambar 2. Sentiment Distribution

Berdasarkan informasi yang ada dalam gambar 2, mayoritas komentar pada konten yang mengandung unsur bahasa isyarat adalah sentimen “Influential”, mencapai sekitar 58.5% dari total komentar atau setara dengan 7.552 komentar. Sementara itu, sentimen “Non- influential” mencakup sekitar 41.5% komentar, atau sebanyak 5.365 komentar. Data ini menunjukkan bahwa banyak masyarakat yang merespon konten tersebut dengan antusias atau tertarik dengan isinya. Sebagai contoh positif, beberapa pengguna Tiktok memberikan dukungan dan semangat kepada pembuat konten, sambil menyatakan bahwa bahasa isyarat adalah hal yang sangat indah. Namun, terdapat pula komentar negatif yang menyoroti perbedaan bahasa isyarat antar negara atau menyerukan untuk tidak terlalu terpengaruh oleh tren konten tersebut. Dengan demikian, kebijakan pembuatan konten yang mengandung unsur bahasa isyarat telah memicu beragam tanggapan dari masyarakat, baik positif, negatif, maupun netral yang menunjukkan minat yang signifikan dari pengguna Tiktok terhadap konten semacam itu.

Word cloud digunakan untuk menggambarkan secara visual frekuensi kemunculan kata-kata dalam dataset, di mana ukuran relatif dari setiap kata dipengaruhi oleh seberapa sering kata tersebut muncul. Sementara itu, distribusi teks memberikan gambaran yang lebih rinci tentang bagaimana kata-kata tersebut tersebar dalam dataset, umumnya diperlihatkan dalam bentuk grafik atau histogram yang menunjukkan distribusi frekuensi kata-kata.



Gambar 3. Non-influential and Influential Word Cloud



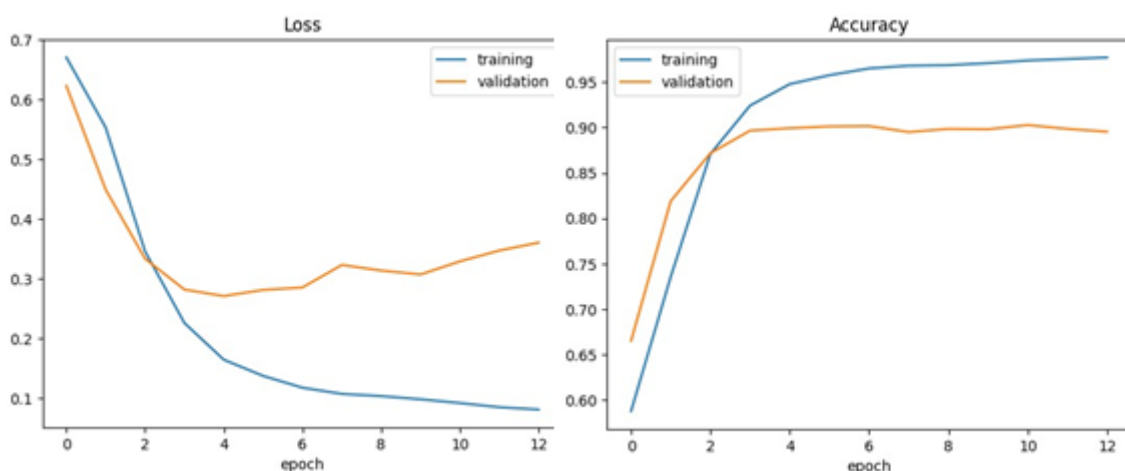
Gambar 4. Distribution of Non-influential and Influential Sentiment Texts

Berdasarkan hasil dari Gambar 3 dan Gambar 4 menunjukkan bahwa ukuran kata dalam word cloud dan distribusi teks meningkat seiring dengan frekuensi kemunculan istilah tersebut dalam analisis. Kedua gambar ini memperlihatkan kata-kata yang paling umum digunakan dalam komentar-komentar pada konten tersebut, dengan ukuran huruf yang lebih besar menandakan frekuensi kata yang lebih tinggi. Namun, terdapat beberapa stopword yang masih terlihat dalam word cloud dan belum dihilangkan. Pada word cloud untuk sentimen Non- influential, masyarakat cenderung lebih banyak membahas topik yang terkandung dalam konten seperti "sign", "language", "twinkling", "watermelon", "deaf", "people", "learn", dan lainnya. Sedangkan dalam word cloud untuk sentimen Influential, masyarakat cenderung lebih menekankan manfaat dari konten tersebut atau menyampaikan perasaan dan semangat kepada pembuat konten, dengan kata-kata seperti "beautiful", "spirit", "love", "exaited", "like", "cry", "hopefully", "thank", dan sebagainya.

3.2 Implementasi dan Evaluasi RNN

Model RNN LSTM digunakan untuk memproses teks yang termasuk dalam analisis sentimen. Arsitektur LSTM dipilih karena kemampuannya memahami konteks temporal dalam teks dan mempertahankan informasi jangka panjang untuk merancang sistem guna menganalisis kredibilitas sentimen yang beragam dan multidimensi dengan cermat. Model ini dilatih menggunakan data teks yang telah dilabeli dengan sentimen yang sesuai seperti Influential dan Non-influential. Dalam penelitian ini, metode Simple Random Sampling diterapkan untuk menentukan data pelatihan dan data pengujian secara acak. Metode ini menjamin bahwa setiap sampel data memiliki peluang yang sama untuk terpilih dalam setiap set data, sehingga menghasilkan representasi yang tidak bias dari keseluruhan populasi data. Sedangkan Prinsip Pareto, yang juga dikenal sebagai aturan 80/20 diterapkan untuk menentukan proporsi pembagian data. Prinsip ini menyatakan bahwa 80% data akan dialokasikan untuk set pelatihan, sedangkan 20% sisanya untuk set pengujian.

Langkah selanjutnya dalam penelitian ini adalah pembuatan model menggunakan arsitektur LSTM. Proses dimulai dengan pembuatan layer embedding sebagai input layer. Layer embedding berperan penting dalam memahami makna kata-kata dalam teks dan mengelompokkan kata-kata yang memiliki makna serupa agar terletak berdekatan dalam ruang vektor. Selanjutnya dilakukan pembuatan layer LSTM dengan jumlah unit sebanyak 128. Layer LSTM bertujuan untuk memproses urutan kata-kata dalam teks dengan mempertahankan informasi jangka panjang. Setelah itu, seluruh jaringan disambungkan pada tiap neuron dengan layer dense yang berjumlah 2 sesuai dengan jumlah kategori pada penelitian. Model dikonfigurasi dengan menggunakan optimasi adam dan categorical cross entropy loss function untuk melakukan pengelompokkan multi-class berdasarkan prediksi yang dihasilkan oleh model.



Gambar 5. RNN Model Loss and Accuracy Graph

Dalam evaluasi performa model, metrik akurasi (acc) mengindikasikan proporsi prediksi yang benar dari data pelatihan, sementara val_acc menunjukkan proporsi prediksi yang benar dari data pengujian. Hasil eksperimen menunjukkan bahwa setelah 13 iterasi pelatihan, akurasi

tertinggi yang dicapai adalah 90%. Pada sisi distribusi kerugian model, loss merujuk pada tingkat kerugian dari data pengujian. Setelah 13 iterasi pelatihan terjadi konvergensi yang mendekati stabilitas, menunjukkan bahwa model telah mencapai tingkat kerugian yang relatif stabil pada kedua set data pelatihan dan pengujian.

Tabel 3. Classification Report RNN LSTM Model

Classification Report	Precision	Recall	f1-score	support
Non-influential	0.90	0.93	0.91	1533
Influential	0.89	0.85	0.87	1048
accuracy			0.90	2581
macro avg	0.89	0.89	0.89	2581
weighted avg	0.90	0.90	0.89	2581

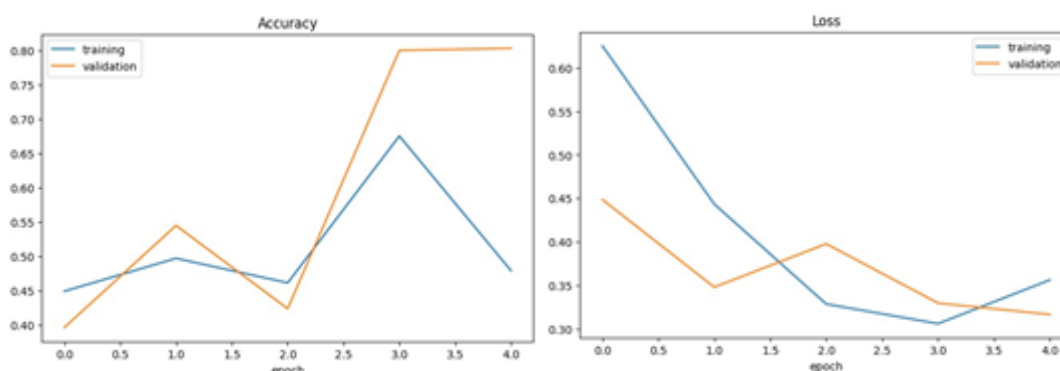
Berdasarkan laporan klasifikasi, model RNN(LSTM) berhasil mencapai tingkat akurasi sebesar 0.90 saat diuji pada data uji. Hasil ini menunjukkan bahwa model tersebut efektif dalam membedakan antara dua kelas, dengan presisi mencapai 0.90% dan recall mencapai 0.93% untuk kelas Influential, serta presisi sebesar 0.89% dan recall sebesar 0.85% untuk kelas Non-influential.

3.3 Implementasi dan Evaluasi BERT

BERT (Bidirectional Encoder Representations from Transformers), merupakan salah satu model representatif dalam pemrosesan bahasa alami yang dikembangkan oleh Google. BERT menonjol dengan memanfaatkan struktur Transformer yang mendasarinya, berbeda dengan pendekatan RNN yang telah digunakan sebelumnya. Transformer merupakan arsitektur yang memungkinkan model untuk memperoleh pemahaman yang lebih mendalam terhadap konteks kata-kata dalam sebuah teks. Salah satu keunggulan BERT adalah kemampuannya untuk memproses teks secara searah maupun dua arah sekaligus.

Secara khusus pelatihan BERT melibatkan dua tahap utama yaitu pra-pelatihan (pre-training) dan penyesuaian (fine-tuning). Proses pra-pelatihan BERT dimulai dengan membagi data menjadi set pelatihan dan pengujian menggunakan metode 'train_test_split' dari library scikit-learn. Setelah menggunakan library Transformers, tokenizer BERT diinisialisasikan dengan memuat model tokenizer dari model BERT yang telah dilatih sebelumnya. Teks pada set pelatihan dan pengujian kemudian di tokenisasi menggunakan tokenizer BERT tersebut.

Setelah teks selesai di-tokenisasi, dataset Tensorflow dibuat dari token-token tersebut dengan memanfaatkan metode 'Dataset.from_tensor_slices()', yang menciptakan dataset TensorFlow dari array tuple tensor. Dalam hal ini, setiap dataset terdiri dari pasangan token-token teks dan label yang sesuai. Model BERT untuk klasifikasi urutan juga diinisialisasi menggunakan metode 'TFBertForSequenceClassification.from_pretrained', dengan menentukan jumlah label yang sesuai dengan jumlah kategori klasifikasi yang diinginkan.



Gambar 6. BERT Model Loss and Accuracy Graph

Pada gambar 6 diatas menunjukkan grafik evaluasi performa model, proses pra- pelatihan BERT dilakukan dengan menentukan optimizer, fungsi kerugian, dan metrik evaluasi yang akan digunakan selama proses pelatihan. Setelah model dikompilasi, proses pelatihan dimulai dengan

memanggil metode 'fit' pada model dengan menggunakan data pelatihan yang telah dipersiapkan sebelumnya. Data pelatihan diacak dengan setiap epoch (iterasi) menggunakan metode 'shuffle' dan dipisahkan menjadi batch dengan ukuran 16. Model dilatih selama 5 epoch dengan batch size sebesar 128. Selama pelatihan, validasi dilakukan menggunakan data pengujian yang telah disiapkan sebelumnya, yang juga dipisahkan menjadi batch dengan ukuran 32. Hasil pelatihan akan disimpan dalam variabel 'hist' yang berisi informasi tentang perubahan metrik evaluasi akurasi selama setiap epoch, sehingga menghasilkan analisis kinerja model secara lebih mendalam setelah pelatihan selesai.

Tabel 4. BERT Model Classification Report

Classification Report	Precision	Recall	f1-score	support
Influential	0.94	0.55	0.70	1048
Non-influential	0.76	0.98	0.86	1534
accuracy			0.80	2582
macro avg	0.85	0.76	0.78	2582
weighted avg	0.83	0.80	0.79	2582

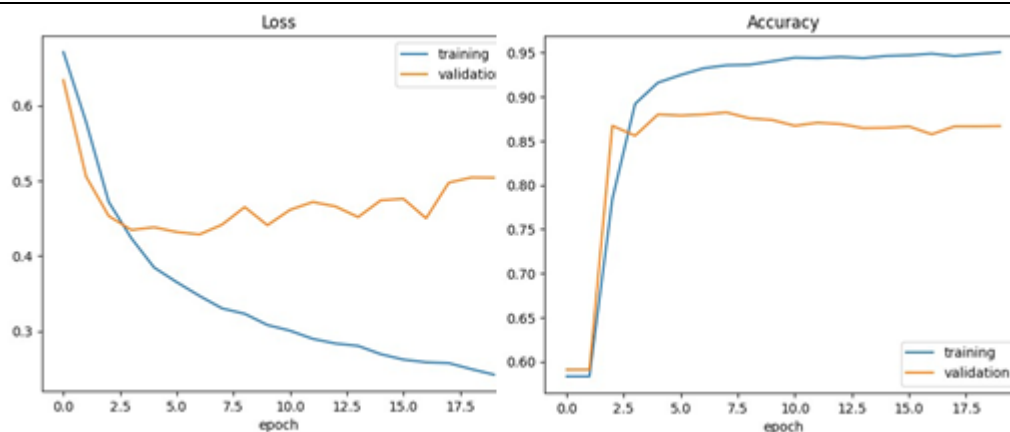
Berdasarkan laporan hasil klasifikasi performa pada Tabel 4, dapat diamati bahwa model BERT mampu mencapai tingkat akurasi sebesar 0.80% yang menunjukkan seberapa baik model dapat mengklasifikasikan sentimen dari data yang diberikan. Ketika meninjau presisi, model menunjukkan tingkat yang baik dalam mengidentifikasi komentar yang tergolong "Influential" dengan nilai presisi sebesar 0.94%. Namun, presisi untuk kategori "Non-influential" sedikit lebih rendah, yaitu 0.76%. Recall, mengukur seberapa baik model dalam menemukan semua instance dari suatu kelas tertentu menunjukkan hasil yang lebih tinggi untuk kategori "Non-influential" yaitu 0.98%, dibandingkan dengan kategori "Influential" yaitu 0.55%.

3.4 Analisa Perbandingan

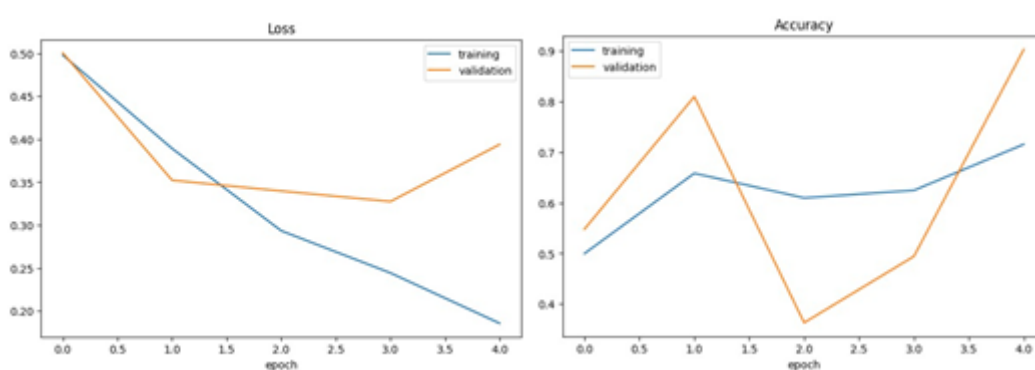
Analisa perbandingan antara model RNN arsitektur LSTM dan model BERT merupakan langkah penting untuk mengevaluasi kemampuan keduanya dalam tugas analisis. Kombinasi antara model RNN dan BERT dilakukan dengan tujuan untuk menggabungkan keunggulan kedua model tersebut dalam analisis sentimen terhadap dataset komentar pada media sosial. Dalam pelatihan model RNN yang dikombinasikan dengan BERT, pendekatan awal melibatkan penggunaan arsitektur LSTM sebagai model RNN.

Proses dimulai dengan pembagian dataset menjadi data latih dan data uji, yang kemudian diikuti dengan pembuatan layer embedding sebagai input layer. Model RNN kemudian dikonfigurasi menggunakan optimasi Adam dan fungsi loss categorical cross-entropy untuk melakukan pengelompokan multi-class berdasarkan prediksi yang dihasilkan. Hasil eksperimen terdokumentasi pada Gambar 5. Selanjutnya, penggabungan antara model RNN dan BERT dilakukan dengan membagi dataset yang sama antara model pertama RNN dan model kedua BERT. Proses ini memungkinkan setiap model untuk menggunakan data yang sama untuk melatih dan menguji performanya, meskipun dengan pendekatan dan metode yang berbeda.

Dalam pelatihan model BERT yang dikombinasikan dengan RNN, proses pra-pelatihan BERT dimulai dengan pembagian dataset menjadi set pelatihan dan pengujian menggunakan metode 'train_test_split' dari pustaka scikit-learn. Selanjutnya, menggunakan pustaka Transformer tokenizer BERT diinisialisasikan dengan memuat model tokenizer yang telah dilatih sebelumnya. Penggabungan antara model BERT dan RNN dilakukan dengan membagi dataset yang sama antara model pertama BERT dan model kedua RNN.



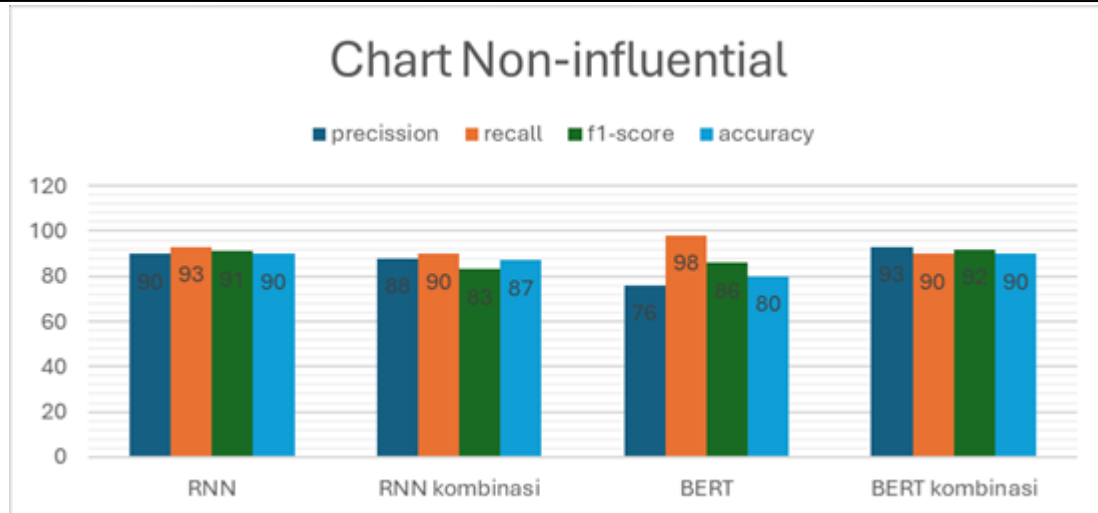
Gambar 7. RNN Loss and Accuracy Graph with BERT Combination



Gambar 8. BERT Loss and Accuracy Graph with RNN Combination

Berdasarkan grafik pada Gambar 7, hasil eksperimen dari model RNN yang dikombinasikan dengan BERT menunjukkan bahwa setelah 15 iterasi pelatihan, akurasi tertinggi yang dicapai adalah 86%. Distribusi kerugian model yang mengacu pada tingkat kerugian dari data pengujian menunjukkan bahwa setelah 15 iterasi pelatihan terjadi konvergensi yang mendekati stabilitas dan model mencapai tingkat kerugian yang relatif stabil pada kedua set data pelatihan dan pengujian. Namun, performa model dalam pengujian RNN yang dikombinasikan dengan BERT mengalami penurunan dibandingkan dengan model sebelumnya yang hanya menggunakan RNN.

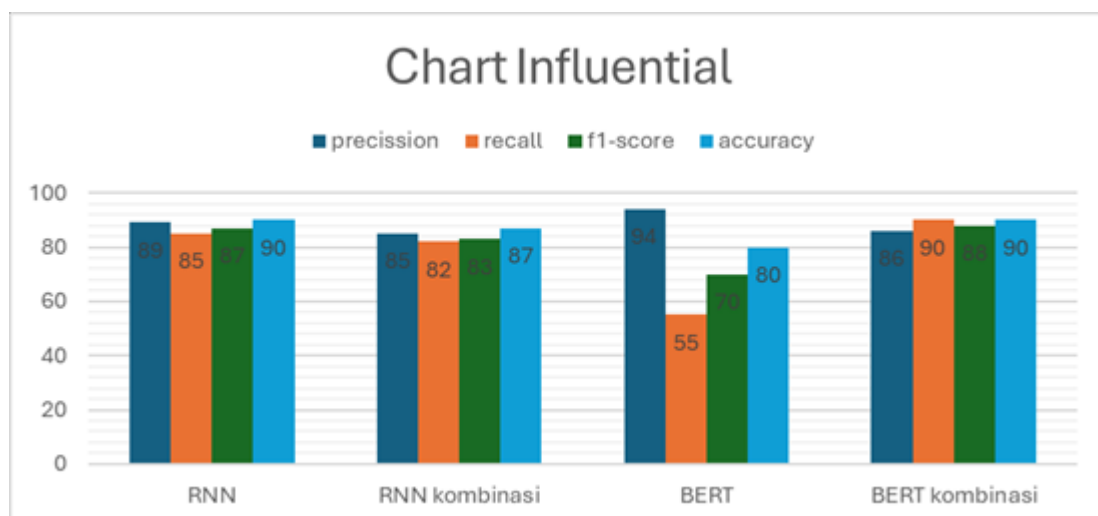
Sementara itu, pada Gambar 8 hasil eksperimen dari model BERT yang dikombinasikan dengan RNN menunjukkan bahwa setelah 5 iterasi pelatihan, akurasi tertinggi yang dicapai adalah 90%. Distribusi kerugian model juga menunjukkan konvergensi yang mendekati stabilitas setelah 5 iterasi pelatihan. Hal ini menandakan bahwa model telah mencapai tingkat kerugian yang relatif rendah pada kedua set data pelatihan dan pengujian. Dalam pengujian BERT yang dikombinasikan dengan RNN, terlihat bahwa performa model mengalami peningkatan dibandingkan dengan model sebelumnya yang hanya menggunakan BERT saja.



Gambar 9. Non-influential Chart

Berdasarkan hasil pengujian pada Gambar 9. Terdapat perbedaan yang signifikan dalam kinerja model yang menggunakan arsitektur RNN, BERT, serta kombinasi keduanya dalam mengklasifikasikan sentimen Non-influential. Model RNN menunjukkan nilai presisi sebesar 90%, recall sebesar 93%, f1-score sebesar 91% dan akurasi sebesar 90%. Model ini memiliki kinerja yang sangat baik dalam mengidentifikasi komentar Non-influential, dengan tingkat presisi dan recall yang tinggi. Namun, model RNN kombinasi menunjukkan penurunan performa dengan nilai presisi sebesar 88%, recall sebesar 90%, f1-score sebesar 83%, dan akurasi sebesar 87%. Meskipun masih menghasilkan hasil yang baik, terlihat bahwa model kombinasi ini tidak sebaik model RNN tunggal dalam mengklasifikasikan sentimen Non-influential.

Di sisi lain, model BERT menunjukkan nilai presisi yang lebih rendah yaitu 76%, namun memiliki recall yang sangat tinggi yaitu 98%, sehingga menciptakan keseimbangan yang baik antara presisi dan recall. Meskipun demikian, nilai f1-score dan akurasi model BERT masih cukup tinggi yaitu 86% dan 80% secara berturut-turut. Ketika digabungkan dengan model RNN, model BERT kombinasi menunjukkan peningkatan yang signifikan dalam nilai presisi menjadi 93%, dengan recall yang relatif stabil di 90%. Hal ini menghasilkan nilai f1-score sebesar 92% dan akurasi 90%, menunjukkan bahwa kombinasi kedua model tersebut menghasilkan kinerja yang lebih baik dalam mengklasifikasikan sentimen Non-influential. Dengan demikian, penggunaan model kombinasi BERT kombinasi RNN dapat menjadi pilihan yang lebih baik dalam analisis sentimen untuk data Non-influential.



Gambar 10. Influential Chart

Berdasarkan hasil pengujian, terlihat perbedaan yang signifikan dalam kinerja model yang digunakan untuk mengklasifikasikan sentimen *Influential*. Model RNN menunjukkan performa yang baik dengan nilai presisi sebesar 89%, recall sebesar 85%, f1-score sebesar 87%, dan akurasi sebesar 90%. Model ini efektif dalam mengidentifikasi komentar yang tergolong sebagai *Influential*, dengan tingkat presisi dan recall yang tinggi. Namun, model RNN kombinasi menunjukkan penurunan performa dengan nilai presisi sebesar 85%, recall sebesar 82%, f1-score sebesar 83%, dan akurasi sebesar 87%. Terlihat bahwa model kombinasi ini tidak sebaik model RNN tunggal dalam mengklasifikasikan sentimen *Influential*.

Di sisi lain, model BERT menunjukkan nilai presisi yang tinggi, yaitu 94%, namun memiliki recall yang rendah yaitu 55%. Hal ini mengindikasikan bahwa model ini cenderung untuk memberikan banyak *false negative* dengan mengabaikan beberapa komentar yang sebenarnya tergolong sebagai *Influential*. Meskipun memiliki nilai f1-score yang cukup rendah yaitu 70% dan akurasi sebesar 80%, model BERT tetap menunjukkan kemampuan yang cukup baik dalam mengklasifikasikan komentar. Ketika digabungkan dengan model RNN, model BERT kombinasi menunjukkan peningkatan yang signifikan dalam nilai recall menjadi 90%, dengan presisi yang relatif stabil di 86%. Hal ini menghasilkan nilai f1-score sebesar 88% dan akurasi 90%, menunjukkan bahwa kombinasi kedua model tersebut menghasilkan kinerja yang lebih baik dalam mengklasifikasikan sentimen *Influential*. Dengan demikian, penggunaan model BERT kombinasi RNN dapat menjadi pilihan yang lebih baik dalam analisis sentimen untuk data *Influential*.

4. Kesimpulan

Berdasarkan analisis sentimen masyarakat terhadap konten bahasa isyarat pada media sosial Tiktok menggunakan model RNN dan BERT, dapat disimpulkan bahwa kedua model tersebut memiliki kelebihan dan kekurangan masing-masing. Model RNN dengan arsitektur LSTM terbukti memiliki performa yang baik dalam mengklasifikasikan sentimen, terutama dalam mengidentifikasi konten yang dianggap sebagai "*Influential*". Hal ini karena RNN mampu menangkap hubungan temporal antar kata dan mempertahankan informasi jangka panjang yang krusial dalam memahami konteks pesan. Kombinasi model RNN dengan BERT menunjukkan peningkatan kinerja dalam mengklasifikasikan sentimen bahwa pendekatan gabungan dapat meningkatkan kemampuan model dalam memahami konten yang kompleks.

Di sisi lain, model BERT menunjukkan kemampuan yang baik dalam memahami konteks kata secara menyeluruh, namun cenderung memiliki kesulitan dalam mengidentifikasi konten yang dianggap "*Influential*", terutama karena kecenderungannya untuk memberikan banyak *false negative*. Oleh karena itu, penggunaan model gabungan RNN dan BERT dapat memberikan wawasan yang lebih baik dalam analisis sentimen untuk konten bahasa isyarat pada media sosial Tiktok, sehingga memungkinkan untuk pemahaman yang lebih komprehensif dan akurat tentang respons pengguna terhadap konten tersebut.

Berdasarkan kesimpulan dari analisis sentimen menggunakan model RNN dan BERT terhadap konten bahasa isyarat pada media sosial Tiktok, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan selanjutnya. Pertama, penting untuk memperluas cakupan dataset dengan mengumpulkan data dari berbagai sumber atau platform media sosial selain Tiktok. Hal ini dapat membantu dalam memperoleh pemahaman yang lebih holistik tentang sentimen masyarakat terhadap bahasa isyarat di berbagai konteks. Selain itu, pengumpulan data dari berbagai negara atau budaya juga dapat memberikan pandangan yang lebih luas dan diversifikasi dalam analisis. Kedua, penting untuk memperhatikan penyesuaian parameter dan teknik preprocessing yang tepat agar model dapat mengatasi tantangan yang spesifik dalam data. Terakhir, penelitian selanjutnya dapat mengeksplorasi penggunaan teknik pemrosesan bahasa alami yang lebih canggih untuk memperoleh representasi teks yang lebih kaya dan kontekstual. Penggunaan teknik seperti word embeddings yang ditingkatkan atau model bahasa berbasis Transformer selain BERT dapat meningkatkan kemampuan model dalam memahami nuansa dan konteks bahasa isyarat.

Referensi

- [1] H. T. Duong and T. A. Nguyen-Thi, "A review: preprocessing techniques and data augmentation for sentiment analysis," *Comput. Soc. Networks*, vol. 8, no. 1, pp. 1–17, 2021, doi: 10.1186/s40649-020-00080-x.

- [2] M. Agrawal, R. Ainapure, S. Agrawal, S. Bhosale, and S. Desai, "Models for Hand Gesture Recognition using Deep Learning," *2020 IEEE 5th Int. Conf. Comput. Commun. Autom. ICCCA 2020*, pp. 589–594, 2020, doi:10.1109/ICCCA49541.2020.9250846.
- [3] H. Utami, "Analisis Sentimen dari Aplikasi Shopee Indonesia Menggunakan Metode Recurrent Neural Network," *Indones. J. Appl. Stat.*, vol. 5, no. 1, p. 31, 2022, doi: 10.13057/ijas.v5i1.56825.
- [4] R. A. Pramunendar, D. P. Prabowo, and R. A. Megantara, "Metode Recurrent Neural Network (Rnn) Dengan Arsitektur Lstm Untuk Analisis Sentimen Opini Publik Terkait Vaksin Covid-19," *J. Inform. Upgris*, vol. 8, no. 1, pp. 44–48, 2022.
- [5] M. J. Islam, R. Datta, and A. Iqbal, "Actual rating calculation of the zoom cloud meetings app using user reviews on google play store with sentiment annotation of BERT and hybridization of RNN and LSTM," *Expert Syst. Appl.*, vol. 223, no. March, 2023, doi: 10.1016/j.eswa.2023.119919.
- [6] S. Siswanto, Z. Mar'ah, A. S. D. Sabir, T. Hidayat, F. A. Adhel, and W. S. Amni, "The Sentiment Analysis Using Naïve Bayes with Lexicon-Based Feature on TikTok Application," *J. Varian*, vol. 6, no. 1, pp. 89–96, 2022, doi: 10.30812/varian.v6i1.2205.
- [7] R. I. Syah, H. Hoiriyah, and M. Walid, "Analisis Sentimen Pengguna Media Sosial Terhadap Aplikasi M-Health Peduli Lindungi Dengan Metode Lexicon Based Dan Naïve Bayes," *Indones. J. Bus. Intell.*, vol. 6, no. 1, 2023, doi: 10.21927/ijubi.v6i1.3275.
- [8] F. Nurdiyansyah and L. U. Pratama, "Analisis sentimen perpindahan ibu kota negara pada aplikasi Tiktok menggunakan metode LSTM," vol. 17, no. September, pp. 382–387, 2023.
- [9] T. H. C. Chiang, C. S. Liao, and W. C. Wang, "Investigating the Difference of Fake News Source Credibility Recognition between ANN and BERT Algorithms in Artificial Intelligence," *Appl. Sci.*, vol. 12, no. 15, 2022, doi: 10.3390/app12157725.
- [10] Y. Yunitasari, A. Musdholifah, and A. K. Sari, "Sarcasm Detection For Sentiment Analysis in Indonesian Tweets," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 13, no. 1, p. 53, 2019, doi: 10.22146/ijccs.41136.
- [11] C. J. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-based Model for," *Eighth Int. AAAI Conf. Weblogs Soc. Media*, pp. 216–225, 2014.
- [12] M. Isnain, G. N. Elwirehardja, and B. Pardamean, "Sentiment Analysis for TikTok Review Using VADER Sentiment and SVM Model," *Procedia Comput. Sci.*, vol. 227, pp. 168–175, 2023, doi: 10.1016/j.procs.2023.10.514.
- [13] S. Lindgren, "Introducing Connected Concept Analysis: A network approach to big text datasets," *Text Talk*, vol. 36, no. 3, pp. 341–362, 2016, doi: 10.1515/text-2016-0016.
- [14] S. Merrill and M. Åkerlund, "Standing up for Sweden? The racist discourses, architectures and affordances of an anti-immigration facebook group," *J. Comput. Commun.*, vol. 23, no. 6, pp. 332–353, 2018, doi: 10.1093/jcmc/zmy018.
- [15] M. Hamka and D. Ratna Sari, "Analisis Sentimen Dan Information Extraction Pembelajaran Daring Menggunakan Pendekatan Lexicon," *Djtechno J. Teknol. Inf.*, vol. 3, no. 1, pp. 21–32, 2022, doi: 10.46576/djtechno.v3i1.2194.
- [16] M. Åkerlund, "The importance of influential users in (re) producing Swedish far-right discourse on Twitter," *Eur. J. Commun.*, vol. 35, no. 6, pp. 613–628, 2020, doi: 10.1177/0267323120940909.
- [17] L. Vu and T. P. D. Le, "A lexicon-based method for Sentiment Analysis using social network data," *Int'l Conf. Inf. Knowl. Eng.*, no. July, p. 146, 2017.
- [18] Mahira Putri, T. E. Sutanto, and S. Inna, "Studi Empiris Model BERT dan DistilBERT Analisis Sentimen pada Pemilihan Presiden Indonesia," *Indones. J. Comput. Sci.*, vol. 12, no. 5, pp. 2972–2980, 2023, doi: 10.33022/ijcs.v12i5.3445.