ANALISIS SENTIMEN ULASAN APLIKASI LINKEDIN BERBASIS LEXICON DAN LONG SHORT-TERM MEMORY (LSTM)

Rizky Putra Kurniawan, Istiadi, Syahroni Wahyu Irianandа
Teknik Informatika, Universitas Widya Gama Malang
Jalan Borobudur No. 35, Mojolangu, Kec. Lowokwaru, Kota Malang, Jawa Timur, Indonesia
rizkyputrakurniawan24@gmail.com

ABSTRAK

Analisis sentimen memiliki peranan yang sangat penting dalam pengolahan data ulasan yang terus berkembang seiring dengan kemajuan dalam teknologi informasi dan komunikasi. Salah satu sumber data yang relevan adalah ulasan aplikasi LinkedIn yang terdapat di *Play Store*, yang mencerminkan pengalaman beragam penggunanya dalam konteks profesional. Mengingat banyaknya jumlah ulasan, pengolahan data secara manual akan memakan waktu yang lama dan beresiko terhadap subjektivitas, sehingga diperlukan metode yang otomatis dan objektif, yaitu analisis sentimen. Penelitian ini bertujuan untuk menganalisis sentimen dari 5.000 ulasan aplikasi LinkedIn dalam periode 2020-2024 dengan pendekatan *Hybrid*, yang menggabungkan pendekatan berbasis *Lexicon* dengan menggunakan InSet Lexicon untuk pelabelan sentimen dan pendekatan *Machine Learning* dengan model *Deep Learning*, yaitu *Long Short-Term Memory* (LSTM) untuk pemodelan. Proses pelabelan sentimen menghasilkan 2.755 ulasan positif (55.1%) dan 2.245 ulasan negatif (44.9%). Sementara itu, proses pemodelan melibatkan pengujian pembagian rasio data dan jumlah epoch, dengan hasil kinerja model terbaik ditemukan pada rasio 50:50 dan 30 epoch, mencapai *accuracy* 88.21%, serta *precision*, *recall*, dan *f1-score* masing-masing sebesar 88%. Temuan ini menunjukkan bahwa pendekatan *Hybrid* yang diterapkan mampu memberikan kinerja analisis sentimen yang handal dan memberikan wawasan berharga mengenai persepsi pengguna terhadap aplikasi LinkedIn.

Kata kunci : *Analisis sentimen, LinkedIn, Pendekatan Hybrid, InSet Lexicon, LSTM Model.*

1. PENDAHULUAN

Perkembangan dalam bidang teknologi informasi dan komunikasi telah mengubah cara individu berinteraksi dan berkomunikasi, terutama melalui platform digital yang memungkinkan penyebaran informasi dengan cepat. Salah satu fenomena yang muncul adalah meningkatnya jumlah ulasan pengguna terhadap layanan digital. Ulasan ini sangat berharga karena dapat memberikan wawasan mengenai persepsi, kepuasan, dan harapan pengguna terhadap layanan yang diberikan. Namun, tantangan yang dihadapi semakin beragam, termasuk kemampuan untuk mengolah dan menganalisis data tersebut [1].

Perkembangan ini mendorong perusahaan untuk memahami lebih dalam mengenai pandangan dan sikap pelanggan terhadap produk atau layanan yang mereka tawarkan [2].

Disisi lain, analisis data secara manual dapat memakan waktu yang lama dan beresiko terpengaruh oleh subjektivitas. Oleh karena itu, diperlukan metode atau teknik yang otomatis dan objektif, yaitu analisis sentimen.

Analisis sentimen merupakan suatu proses yang sistematis dalam mengidentifikasi dan meng-kategorikan perasaan atau opini individu terhadap berbagai layanan, seperti film, aplikasi digital, dan produk ke dalam tiga kategori: positif, negatif, atau netral [3].

Dalam pelaksanaan analisis sentimen, terdapat tiga pendekatan utama yang umum digunakan, yaitu pendekatan berbasis *Lexicon*, pendekatan *Machine Learning*, dan pendekatan *Hybrid* [4].

Pada penelitian ini mengadopsi pendekatan *Hybrid*, yang mengintegrasikan dua pendekatan untuk menyelesaikan suatu masalah dengan memanfaatkan keunggulan masing-masing pendekatan, yaitu pendekatan berbasis *Lexicon* dan pendekatan *Machine Learning* [5].

Analisis sentimen dengan pendekatan *Hybrid* ini diterapkan pada data ulasan aplikasi LinkedIn berbahasa Indonesia yang dikumpulkan dari tahun 2020 hingga 2024, dengan total 5.000 ulasan yang tersedia di *Play Store*. Data yang digunakan sudah dilengkapi dengan *rating* yang mencerminkan tingkat kepuasan atau pandangan pengguna terhadap aplikasi. Namun, sering kali *rating* tersebut tidak merepresentasikan keseluruhan pandangan pengguna. Penulis mengamati adanya beberapa ulasan atau pandangan yang tidak sejalan dengan *rating* yang diberikan. Oleh karena itu, diperlukan penerapan metode atau teknik analisis sentimen dengan pendekatan *Hybrid* ini, dengan harapan dapat memperoleh hasil yang lebih baik dan akurat dibandingkan hanya mengandalkan *rating*.

Tujuan dari penelitian ini adalah menerapkan pendekatan *Hybrid* untuk proses analisis sentimen serta untuk mengatasi keterbatasan yang ada pada pendekatan berbasis *Lexicon* dan pendekatan *Machine Learning* terutama model tradisional dalam hal kapasitas untuk memahami konteks yang lebih kompleks dalam data ulasan [6].

Untuk mencapai tujuan tersebut, penelitian ini menerapkan kombinasi antara pendekatan berbasis *Lexicon* menggunakan kamus atau sentimen leksikon

yaitu Indonesian Sentiment (InSet) Lexion dan pendekatan *Machine Learning* terutama model *Deep Learning*, yaitu *Long Short-Term Memory* (LSTM) dalam proses analisis sentimen berbahasa Indonesia pada ulasan aplikasi LinkedIn yang tersedia di *Play Store*. Kombinasi ini memberikan keuntungan yang berbeda dalam analisis sentimen dengan memanfaatkan keunggulan masing-masing pendekatan dalam menangani data teks. Pendekatan berbasis *Lexicon* memungkinkan identifikasi sentimen yang cepat melalui penggunaan kamus kata-kata yang telah diberi nilai polaritas sesuai dengan sentimennya [7].

Namun, pendekatan ini memiliki keterbatasan dalam memahami konteks kalimat yang lebih rumit. Untuk mengatasi masalah ini, pendekatan *Machine Learning*, khususnya model *Deep Learning* yaitu LSTM digunakan untuk memperkuat proses analisisnya. Kemampuan model LSTM dalam membedakan hubungan temporal dan nuansa kontekstual dalam teks dengan mempertimbangkan urutan kata [8] menjadikannya pilihan yang ideal untuk mengatasi keterbatasan tersebut.

2. TINJAUAN PUSTAKA

Pada bagian ini berisi hasil penelitian terdahulu dan tinjauan pustaka yang mencakup teori-teori dan konsep yang relevan dengan penelitian ini.

2.1. Penelitian Terdahulu

Banyak penelitian telah dilakukan mengenai ulasan aplikasi LinkedIn dalam bahasa Indonesia. Namun, mayoritas penelitian tersebut mengadopsi model *Traditional Machine Learning*. Contohnya, penelitian yang dilakukan oleh Al-Husna et al. [9] yang membandingkan performa model *Support Vector Machine* (SVM) dan *Naïve Bayes* pada dataset yang terdiri dari 2.000 ulasan aplikasi LinkedIn. Temuan dari penelitian ini menunjukkan bahwa model SVM memberikan kinerja yang lebih baik dalam analisis sentimen dengan tingkat *accuracy* mencapai 89%, serta *precision*, *recall*, dan *f1-score* masing-masing sebesar 88%. Namun, proses pelabelan sentimen dalam penelitian ini dilakukan secara manual, yang dimana merupakan kelemahan signifikan karena memakan waktu dan beresiko terpengaruh oleh subjektivitas. Penelitian ini juga merekomendasikan untuk mengintegrasikan model *Deep Learning*, seperti LSTM, guna menangani ulasan yang lebih kompleks. Penelitian terdahulu selanjutnya juga telah dilakukan oleh Defriani et al. [10] menunjukkan bahwa metode berbasis *Lexicon* memiliki keterbatasan dalam memahami konteks kalimat dalam ulasan, serta ketergantungan pada sentimen leksikon yang digunakan. Penelitian ini menggunakan dataset yang terdiri dari 10.000 ulasan aplikasi LinikedIn yang dikumpulkan antara 12 September 2018 hingga 20 September 2022. Namun, setelah melalui beberapa tahap praproses teks, jumlah ulasan menyusut menjadi 6.821 ulasan. Sentimen leksikon yang diterapkan dalam penelitian ini adalah VADER Lexicon, yang menghasilkan temuan sebagai berikut: 3.432 ulasan (56.56%) memilki sentimen positif, 1.875 ulasan (20.62%) netral, dan 779 ulasan (12.82%) negatif. Namun, karena VADER Lexicon memerlukan teks dalam bahasa Inggris untuk analisis, ulasan yang ditulis dalam bahasa Indonesia harus diterjemahkan terlebih dahulu agar dapat dianalisis oleh metode tersebut. Selain itu, penelitian selanjutnya yang menggunakakn model *Traditional Machine Learning*, khususnya SVM, dilakukan oleh Irianandа et al. [11].

Dalam penelitiannya, dilakukan pengoptimalan klasifikasi sentimen pada komentar pengguna *mobile game*. Untuk mencapai tujuan tersebut, kombinasi SVM, Grid Search, dan N-Gram digunakan, yang menghasilkan nilai *accuracy* tertinggi sebesar 87.33%. Oleh karena itu, dalam penelitian ini yang mengadopsi pendekatan *Hybrid* dalam proses analisis sentimen terhadap dataset ulasan aplikasi LinkedIn berbahasa Indonesia yang tersedia di *Play Store*, diharapkan dapat mengatasi berbagai keterbatasan yang terdapat pada penelitian sebelumnya dan memberikan kontribusi dalam penerapan teknik atau metode baru untuk proses analisis sentimen.

2.2. Analisis Sentimen

Analisis sentimen merupakan suatu proses yang melibatkan pengumpulan serta penelaahan pandangan, pemikiran, dan persepsi masyarakat terkait berbagai isu, produk, subjek, dan layanan. Proses analisis ini sangat berguna bagi perusahaan, instansi pemerintahan, maupun individu dalam mengumpulkan informasi dan mengambil keputusan yang didasarkan pada informasi yang diperoleh dari analisis tersebut [12].

2.3. Pendekatan Analisis Sentimen

Pendekatan berbasis *Lexicon* memanfaatkan sentimen leksikon untuk mengidentifikasi dan menentukan orientasi sentimen. Sentimen leksikon merupakan kumpulan kata atau kamus yang telah diberi label sesuai dengan orientasi semantiknya [13].

Pendekatan ini sangat bergantung pada kelengkapan dan kualitas sentimen leksikon yang digunakan, sehingga hasil analisis dapat bervariasi tergantung pada domain atau bahasa yang diteliti. Selain itu, pendekatan ini relatif mudah untuk diterapkan dan memberikan interpretasi yang intuitif. Di sisi lain, pendekatan *Machine Learning* dalam analisis sentimen terbagi menjadi dua kategori, yaitu *Traditional Machine Learning Model* seperti *Support Vector Machine* (SVM), *Naïve Bayes*, atau *Maximum Entropy*, serta *Deep Learning Model* seperti *Convolutional Neural Network* (CNN) dan *Recurrent Neural Network* (RNN) [14].

Sedangkan pendekatan *Hybrid* adalah suatu metode yang mengintegrasikan dua pendekatan untuk mengatasi suatu masalah dengan memanfaatkan keunggulan dan masing-masing pendekatan. Dalam konteks ini, kombinasi pendekatan yang digunakan

adalah pendekatan berbasis *Lexicon* dan pendekatan *Machine Learning.*

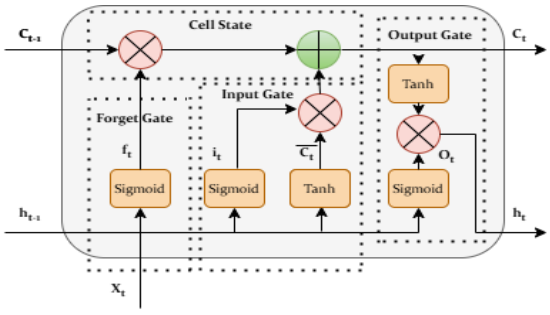
2.4. Indonesian Sentiment (InSet) Lexicon

Indonesian Sentiment (InSet) Lexicon adalah hasil penelitian yang dilakukan oleh Fajri Koto dan Gemala Y. Rahmaningtyas pada tahun 2017. Dalam penelitian tersebut, mereka mengembangkan InSet Lexicon yang mencakup 3.609 kata positif dan 6.069 kata negatif dalam bahasa Indonesia. Setiap kata telah diberi nilai polaritas secara manual berdasarkan sentimennya, dan proses ini ditingkatkan dengan penambahan *stemming* serta penetapan sinonim. Nilai polaritas dalam sentimen leksikon ini berkisar antara -5 (sangat negatif) hingga +5 (sangat positif) [15].

2.5. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) dirancang untuk mengatasi kelemahan yang ada pada *Recurrent Neural Network* (RNN). Salah satu kelemahan tersebut adalah kapasitas memori jangka panjang yang terbatas, yang menghambat kemampuan untuk menyimpan informasi dalam urutan teks yang lebih panjang [16].

Oleh karena itu, model LSTM dipilih dalam penelitian ini karena keunggulannya yang telah terbukti dalam memproses data sekuensial seperti teks, terutama ketika diperlukan konteks jangka panjang. LSTM telah banyak diterapkan dalam *Speech Recognition*, seperti yang ditunjukkan dalam penelitian yang dilakukan oleh Riyanto et al. [17], serta dalam *Emotional Analysis*, dan *Text Analysis*, berkat kemampuannya dalam menyimpan memori dan membuat prediksi yang cukup akurat. Untuk memberikan pemahaman yang lebih mendalam, Gambar 1 berikut ini menggambarkan konfigurasi arsitektur unit LSTM yang mencakup pemrosesan data melalui *gates*, pembaruan *cell state*, dan keluaran hasil melalui *hidden states.*



Gambar 1. LSTM unit

Tahap awal dalam konstruksi model LSTM adalah pembentukan *forget gate*. Langkah ini sangat penting, karena tidak semua informasi dari *timestamp* sebelumnya akan selalu relevan dengan langkah berikutnya. Rumus *forget gate* direpresentasikan pada persamaan 1 berikut:

$$f_t = \sigma\left(W_f\left[h_{t-1,}\, X_t\right] + b_f\right) \quad (1)$$

Forget gate bertanggung jawab untuk menentukan informasi mana dari *timestamp* sebelumnya yang harus dibuang. *Forget gate* dihitung menggunakan fungsi Sigmoid (σ), yang menghasilkan nilai antara 0 dan 1, yang menunjukkan sejauh mana informasi dilupakan. Perhitungan tersebut memerlukan penjumlahan tertimbang dari *hidden state* sebelumnya (h_{t-1}), *input* saat ini (X_t), dan bias (b_f). Matriks bobot (W_f) berfungsi untuk menentukan kepentingan relatif dari masukan ini dalam perhitungan.

Selanjutnya, *input gate* berfungsi untuk menentukan informasi spesifik yang akan dimasukkan ke dalam *cell state. Input gate* terdiri dari dua tahap utama: pertama, pemilihan informasi yang akan diperbarui, kedua, pembuatan kandidat pembaruan untuk dimasukkan ke dalam *cell state*. Dua rumus digunakan dalam *input gate*, seperti yang dapat dilihat pada persamaan 2 dan 3 di bawah ini:

$$i_t = \sigma(W_i \,.\, [h_{t-1}, X_t] + b_i) \quad (2)$$

$$\hat{C}_t = \tanh(W_C \,.\, [h_{t-1} + X_t] + b_C \quad (3)$$

Input gate merupakan gerbang kedua dalam LSTM unit, dan fungsinya adalah untuk menentukan informasi mana yang akan disimpan dalam *cell state.* Gerbang ini terdiri dari lapisan Sigmoid dan lapisan tanh. Seperti yang diuraikan pada persamaan 2, lapisan sigmoid menentukan nilai mana yang akan diperbarui. Lapisan tanh menghasilkan nilai baru $\hat{C}_t$, seperti yang diilustrasikan dalam persamaan 3.

Pada tahap akhir, *output gate* bertugas mengidentifikasi bagian *cell state* yang akan ditransmisikan sebagai *hidden State* untuk *timestamp* berikutnya. Persamaan untuk *output gate* dapat dilihat dalam persamaan 4 dan 5 di bawah ini:

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o) \quad (4)$$

$$h_t = O_t \tanh(C_t) \quad (5)$$

Rumus *output gate*, sebagaimana direpresentasikan pada persamaan 4, bertanggung jawab untuk menentukan informasi mana dari *cell state* yang akan diteruskan ke keluaran akhir unit LSTM pada *timestamp* tertentu. Nilai O_t diperoleh melalui penerapan fungsi aktivasi Sigmoid (σ) pada bobot (W_o) dengan kombinasi *hidden state* saat ini (h_t) dan *cell state* yang telah diaktifkan oleh fungsi tanh (C_t). Selama proses pelatihan, bobot (W_o) diperbarui secara terus-menerus untuk mengoptimalkan kinerja model LSTM. Bias (b_o) berfungsi untuk menggeser *output* fungsi aktivasi. Hasil akhir dari perhitungan ini adalah nilai O_t, sebagaimana didefinisikan dalam persamaan 5. Nilai ini kemudian digunakan dalam perhitungan *output* akhir (h_t) unit LSTM.

2.6. Evaluasi

Metrik evaluasi memiliki peranan yang sangat krusial dalam mencapai model yang optimal. Oleh karena itu, pemilihan metrik evaluasi yang tepat menjadi faktor kunci dalam memperoleh model pengklasifikasian yang terbaik. Secara umum, banyak model yang menggunakan *accuracy* sebagai ukuran untuk menilai kinerja model. Namun, *accuracy* memiliki beberapa kelemahan, yaitu kurang

memberikan informasi yang mendalam dan cenderung bias terhadap data dari kelas mayoritas [18].

Untuk masalah klasifikasi biner yang diteliti dalam penelitian ini, metrik evaluasi dapat didefinisikan berdasarkan *confusion matrix* seperti yang ditampilkan pada Tabel 1 berikut.

Tabel 1. Confusion matrix

		Actual Class	
		Negative (N)	Positive (P)
Predicted Class	Negative (N)	True Negative (TP)	False Positive (FP)
	Positive (P)	False Negative (FN)	True Positive (TP)

Secara umum, *confusion matrix* berfungsi untuk menganalisis kinerja model dengan cara menghitung jumlah kesalahan dan prediksi benar dari model pada data pengujian, serta menghitung metrik evaluasi lainnya, seperti *accuracy, precision, recall,* dan *f1-score.*

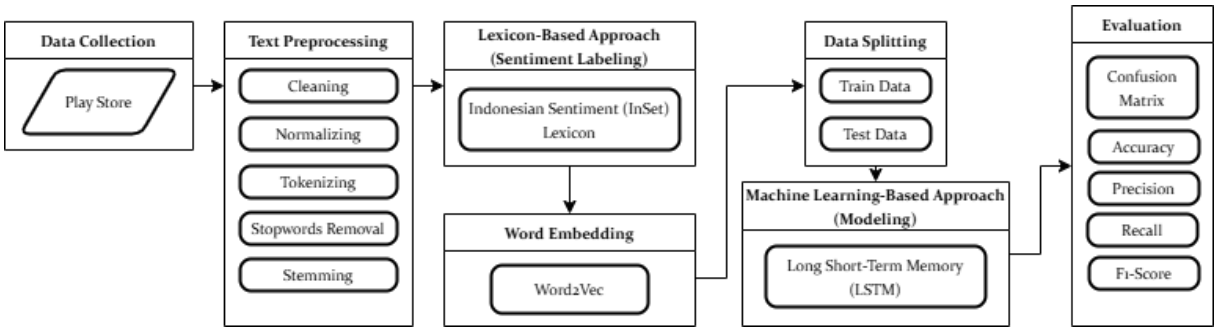
2.7. LinkedIn

LinkedIn adalah salah satu aplikasi jejaring profesional yang paling banyak digunakan di seluruh dunia, dengan basis pengguna yang mencakup pencari kerja, perekrut, dan pemberi kerja [19].

Banyaknya ulasan yang tersedia, menawarkan contoh representatif pengalaman pengguna yang autentik dari berbagai latar belakang. Keragaman pengalaman ini menyediakan kumpulan data yang lengkap yang dapat digunakan untuk menganalisis kualitas layanan, fitur, dan tingkat kepuasan pengguna aplikasi LinkedIn.

3. METODE PENELITIAN

Metode penelitian yang diusulkan dalam penelitian ini terdiri dari beberapa langkah, seperti yang ditunjukkan pada Gambar 2 di bawah ini. Serta penjelasan rinci tentang tahap-tahap penelitian diberikan dalam subbab 3.1 sampai 3.7.



Gambar 2. Tahapan metode penelitian

3.1. Data

Data yang digunakan dalam penelitian ini merupakan data ulasan berbahasa Indonesia dari aplikasi LinkedIn yang terdapat pada *Play Store* yang diambil dari tahun 2020 hingga 2024. Pengumpulan data dilakukan dengan teknik *scraping data* menggunakan *Application Programming Interface* (API) yang disediakan oleh *Google Play Store.* Proses *scraping data* dilaksanakan dengan memanfaatkan *library* Python yang dikenal sebagai *google-play-scraper.*

3.2. Text Preprocessing

Mengacu pada fakta bahwa data yang digunakan dalam penelitian ini berasal dari platform media sosial, khususnya *Play Store*, sangatlah krusial untuk menerapkan teknik *text preprocessing* guna memastikan kualitas dan normalisasi data sebelum analisis dilakukan. Hal ini disebabkan oleh keberagaman bentuk teks yang terdapat di media sosial, seperti singkatan, emotikon, bahasa gaul, angka, dan kalimat yang tidak utuh [20].

Pada penelitian ini, tahapan pertama dalam *text preprocessing* adalah *cleaning.* Pada tahap ini, data mentah akan diperbaiki dengan menghapus karakter atau simbol yang tidak perlu, termasuk emoji, tagar, angka, tanda baca, dan karakter khusus. Tujuan dari tahapan awal ini adalah untuk mengatasi ketidakkonsistenan dalam teks [21].

Tahapan selanjutnya adalah *normalization*, yang mencakup pengelolaan istilah *slang* yang terdapat dalam teks. Tujuan dari tahapan ini adalah untuk menyederhanakan teks dengan menyelaraskan teks dari elemen-elemen yang berbeda agar memiliki makna yang konsisten [22].

Tahapan berikutnya adalah *tokenization*, yang melibatkan pemecahan teks menjadi kata-kata atau token individual untuk memudahkan pemrosesan selanjutnya. Dalam linguistik, *tokenization* merujuk pada metode segmentasi kalimat menjadi kata-kata

individual, yang kemudian disebut sebagai "token" [23].

Setelah teks telah menjadi kata individual atau token, tahapan selanjutnya adalah *stopword removal*. Ini mencakup penghilangan kata-kata umum yang tidak memiliki makna signifikan dalam analisis sentimen, seperti "dan", "kemudian", atau "dengan". Tahapan terakhir yaitu *stemming*, yaitu proses yang mengubah kata menjadi bentuk dasarnya [24].

Dalam penelitian ini, digunakan Sastrawi, sebuah pustaka Python yang dirancang untuk bahasa Indonesia, untuk keperluan *stemming.*

3.3. Labeling

Proses pelabelan dengan InSet Lexicon mencakup perhitungan polaritas sentimen teks melalui penjumlahan nilai polaritas yang terkait dengan setiap kata. Skor polaritas dialokasikan untuk setiap kata sesuai denga kemunculannya dalam sentimen leksikon yang telah ditentukan sebelumnya. Skor sentimen keseluruhan teks ditentukan oleh persamaan 6 berikut:

$$Sentiment\ Score = \sum_{i=1}^{n} Polarity\ (w_i) \quad (6)$$

Dalam konteks ini, variabel "*n*" mewakili jumlah total kata dalam teks, sedangkan polaritas (w_i) menunjukkan skor polaritas yang terkait dengan kata ke-1. Skor sentimen yang melebihi 0 menunjukkan polaritas positif teks. Sebaliknya, jika skor kurang dari 0, teks tersebut diklasifikasikan sebagai negatif.

3.4. Word Embedding

Dalam penelitian ini, teknik Word2Vec digunakan untuk tujuan *word embedding.* Word2Vec merupakan metode *Natural Language Processing* (NLP) yang mengubah korpus teks menjadi vektor numerik, dengan setiap kata direpresentasikan sebagai vektor. Salah satu fitur Word2Vec yang paling signifikan adalah kemampuannya menangkap hubungan semantik dan makna kontekstual. Kemampuan ini berasal dari representasi vektor, yang mengkodekan hubungan antar kata berdasarkan kesamaan kontekstual-nya [25].

3.5. Data Spliting

Dataset penelitian terdiri dari dua komponen utama: data pelatihan dan data pengujian. Data pelatihan digunakan untuk melatih model, sehingga model akan mengenali pola, hubungan, dan struktur dari data tersebut. Data pengujian berfungsi untuk mengevaluasi kinerja model pada data baru yang belum pernah dilihat sebelumnya. Proporsi data pelatihan dan data pengujian merupakan elemen penting yang mempengaruhi efektivitas model, seperti yang ditunjukkan dalam referensi [26].

Tabel 2 berikut mengilustrasikan distribusi proporsi antara data pelatihan dan pengujian yang dilakukan pada penelitian ini.

Tabel 2. Rasio data split

Training Data	Testing Data
70 %	30 %
60 %	40 %
50 %	50 %
40 %	60 %
30 %	70 %

Rasio pembagian data ini dipilih berdasarkan praktik umum dalam evaluasi model *Machine Learning,* seperti yang dibahas dalam [27].

3.6. Pemodelan LSTM

Data yang telah terlabeli menggunakan InSet Lexicon, kemudian data tersebut akan dilakukan proses pemodelan dengan algoritma LSTM. Berikut merupakan model arsitektur dari algoritma LSTM yang digunakan pada penelitian ini.

Tabel 3. Arsitektur model LSTM

Parameter	Value
Neuron	128
Layer	1
Optimizer	Adam
Fungsi Aktivasi	Sigmoid
Loss Function	Binary Crossentropy
Batch Size	64
Epoch	[10, 20, 30]

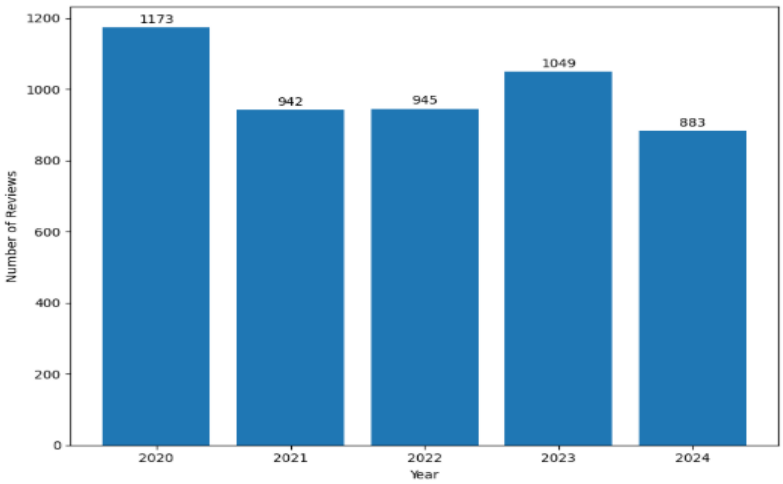
Berdasarkan Tabel 3 di atas, diketahui model arsitektur LSTM yang digunakan pada penelitian ini. Akan tetapi, pada penelitian ini, penulis melakukan *hyperparameter tuning* yang hanya berfokus pada jumlah *epochs,* yang memiliki dampak langsung pada performa model tanpa mengubah struktur atau kompleksitas arsitektur model.

3.7. Evaluasi

Hasil dari proses pemodelan, kemudian akan dilakukan evaluasi untuk menilai kinerja model terhadap data yang belum pernah dilihat oleh model. Tujuannya adalah untuk mengetahui seberapa baik model dapat memahami pola dari data pelatihan dan menerapkannya pada data pengujian. Pemilihan evaluasi metrik pada penelitian ini adalah *accuracy, precision, recall,* dan *f1-score.*

4. HASIL DAN PEMBAHASAN

Hasil proses *scraping data* menghasilkan total 5.000 data ulasan aplikasi LinkedIn di *Play Store* dan sebaran data dari tahun 2020 sampai dengan tahun 2024 dapat dilihat pada Gambar 3 berikut.



Gambar 3. Distribusi ulasan aplikasi LinkedIn tahun 2020-2024

Distribusi jumlah ulasan aplikasi LinkedIn di *Play Store* dari tahun 2020 hingga 2024, seperti yang ditunjukkan pada Gambar 3, menghasilkan total 5.00 ulasan. Penting untuk dicatat bahwa data yang diperoleh hanya tersedia dalam bahasa Indonesia. Dari analisis tersebut, terlihat bahwa pada tahun 2020 tercatat jumlah dengan ulasan tertinggi, yaitu 1.173 ulasan, sementara pada tahun 2024 tercatat jumlah ulasan terendah dengan 883 ulasan. Lonjakan jumlah ulasan pada tahun 2020 dapat diartikan sebagai tanda meningkatnya minat pengguna atau adanya perubahan signifikan pada aplikasi LinkedIn yang mendorong pengguna untuk memberikan ulasan. Sebaliknya, penurunan jumlah ulasan pada tahun 2024 mungkin disebabkan oleh fakta bahwa data tersebut diambil pada pertengahan tahun.

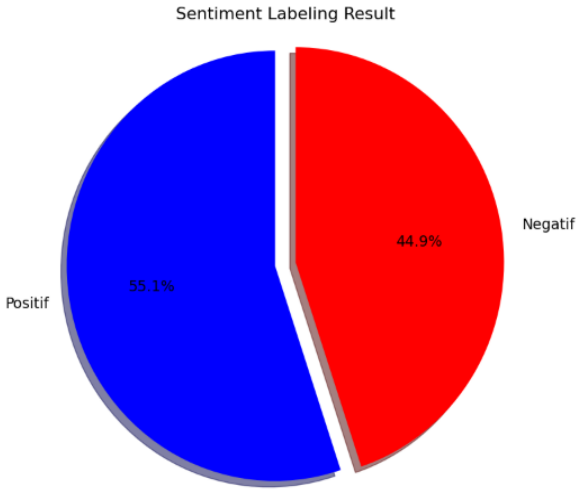
Setelah data berhasil dikumpulkan, tahapan-tahapan *text preprocessing* akan di mulai. Hasil dari sampel data ulasan yang telah melalui semua tahapan *text preprocessing* disajikan dalam Tabel 4 di bawah ini.

Tabel 4. Sampel hasil ulasan hasil text preprocessing

Sebelum	Sesudah
Ketika pendaftaran, verifikasi selalu salah gambar yang manapun dipilih tetap bakal salah, mohon perbaikannya 🙏	daftar verifikasi salah gambar mana pilih salah mohon baik

Berdasarkan Tabel 4 yang ditampilkan, sampel teks telah melalui serangkaian tahapan *text preprocessing*, dan menghasilkan teks yang lebih terstruktur dan berkualitas untuk digunakan dalam langkah-langkah berikutnya. Terlihat bahwa pada sampel teks tersebut telah melewati tahapan penghapusan tanda baca, emoji, serta pemulihan kata ke bentuk baku-nya.

Selanjutnya, data ulasan yang telah terstruktur akan diberi label menggunakan pendekatan berbasis *Lexicon* dengan sentimen leksikon InSet. Hasil dari proses pelabelan untuk seluruh data ulasan ditampilkan pada Gambar 4 berikut.



Gambar 4. Hasil pelabelan InSet Lexicon

Hasil dari proses pelabelan terhadap seluruh data ulasan yang berjumlah total 5.000 menunjukkan bahwa terdapat 2.755 ulasan (55.1%) yang mencerminkan sentimen positif, sementara 2.245 ulasan (44.9%) mencerminkan sentimen negatif. Temuan ini mengindikasikan bahwa selama periode antara tahun 2020 hingga 2024, mayoritas ulasan yang diberikan untuk aplikasi LinkedIn di *Play Store* memiliki sentimen positif, meskipun selisih antara sentimen positif dan negatif tergolong kecil.

Setelah data terlabeli, langkah selanjutnya adalah menerapkannya dalam proses pemodelan dengan menggunakan LSTM model. Tabel 5, 6, 7, 8, dan 9 berikut akan menyajikan gambaran komprehensif mengenai hasil pemodelan, dengan fokus khusus pada berbagai percobaan yang dilakukan dalam penelitian ini, termasuk percobaan pembagian rasio data dan jumlah epoch untuk mendapatkan kinerja model yang terbaik.

Tabel 5. Hasil kinerja model pada pembagian rasio data 70:30

Epochs	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score
10	725	584	109	69	88.03 %	88 %	88 %	88 %
20	719	593	100	75	88.23 %	88 %	88 %	88 %
30	691	613	80	103	87.69 %	88 %	88 %	88 %

Hasil kinerja model pada pembagian rasio 70:30 untuk berbagai jumlah epoch (10, 20, dan 30) ditunjukkan dalam Tabel 5. Dari hasil yang diperoleh, model menunjukkan *accuracy* tertinggi sebesar 88.23% pada epoch ke-20, dengan nilai *precision*, *recall*, dan *f1-score* yang tetap konsisten untuk semua epoch. Secara keseluruhan, performa model pada pembagian rasio data ini menunjukkan stabilitas pada metrik-metrik tersebut, meskipun terdapat fluktuasi kecil pada jumlah TP, TN, FP, dan FN di setiap epoch.

Tabel 6. Hasil kinerja model pada pembagian rasio data 60:40

Epochs	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score
10	935	777	134	136	86.38 %	86 %	86 %	86 %
20	1003	739	172	68	87.39 %	88 %	87 %	88 %
30	1027	694	217	44	86.63 %	88 %	86 %	86 %

Tabel 6 di atas menunjukkan hasil kinerja model berdasarkan pembagian data dengan rasio 60:40 untuk tiga jumlah epoch yang berbeda. Dari hasil yang disajikan, dapat dilihat bahwa *accuracy* tertinggi diperoleh pada epoch ke-20, yaitu sebesar 87.39%, dengan nilai *precision* maksimum mencapai 88% dan *recall* sebesar 87%. Pada epoch ke-30, terdapat sedikit penurunan *accuracy* menjadi 86.63%, meskipun *precision* tetap tinggi pada angka 88%. Hal ini menunjukkan bahwa model tetap menunjukkan performa yang konsisten dalalm menganalisis data, meskipun terdapat variasi kecil dalam distribusi hasil TP, TN, FP, dan FN di setiap epoch.

Tabel 7. Hasil kinerja model pada pembagian rasio data 50:50

Epochs	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score
10	1275	892	219	91	87.48 %	88 %	87 %	87 %
20	1276	902	209	90	87.93 %	88 %	87 %	88 %
30	1240	945	166	126	88.21 %	88 %	88 %	88 %

Tabel 7 di atas menunjukkan hasil kinerja model pada data dengan pembagian rasio 50:50 menggunakan tiga variasi epoch. Berdasarkan tabel yang disajikan *accuracy* tertinggi sebesar 88.21% dicapai pada epoch ke-30, sementara *accuracy* pada epoch ke-20 sedikit lebih rendah, yaitu 87.93%. Model ini juga menunjukkan nilai *precision* yang stabil sebesar 88% di semua epoch. Selain itu, nilai *recall* dan *f1-score* tertinggi, yang juga mencapai 88%, diperoleh pada epoch ke-30, menunjukkan bahwa performa model semakin optimal dalam mendeteksi kelas secara keseluruhan.

Tabel 8. Hasil kinerja model pada pembagian rasio data 40:60

Epochs	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score
10	1562	995	356	60	86.01 %	88 %	85 %	85 %
20	1458	1137	214	164	87.29 %	87 %	87 %	87 %
30	1467	1139	212	155	87.66 %	88 %	87 %	88 %

Berdasarkan Tabel 8 di atas, hasil kinerja model pada data dengan pembagian rasio 40:60 ditunjukkan. Model diuji dengan tiga variasi epoch, yang menunjukkan peningkatan kinerja seiring dengan bertambahnya epoch. Pada epoch ke-30, model mencapai tingkat *accuracy* tertinggi sebesar 87.66%, dengan nilai *precision* dan *recall* yang stabil pada angka 88% dan 87%, serta *f1-score* sebesar 88%. Hal ini mengindikasikan bahwa model dapat mempertahankan keseimbangan dalam mendeteksi kelas positif dan negatif meskipun terdapat perbedaan dalam pembagian data.

Tabel 9. Hasil kinerja model pada pembagian rasio data 30:70

Epochs	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-Score
10	1646	1313	271	238	85.32 %	85 %	85 %	85 %
20	1754	1285	299	130	87.63 %	88 %	87 %	87 %
30	1764	1286	298	120	87.95 %	89 %	87 %	88 %

Tabel 9 di atas menampilkan hasil kinerja model pada data dengan pembagian rasio 30:70. Kinerja model juga diuji dengan tiga variasi epoch yang sama seperti sebelumnya. Model mencapai tingkat *accuracy* tertinggi sebesar 87.63% pada epoch ke-30. Selain itu nilai *precision* dan *recall* masing-masing tercatat pada angka 88% dan 87%, dengan *f1-score* mencapai 88%. Hasil ini menunjukkan bahwa model tetap konsisten

dalam mendeteksi performa optimal meskipun jumlah data latih lebih sedikit dibandingkan dengan data uji.

Hasil dari beberapa percobaan yang dilakukan dengan beberapa rasio pembagian data dan jumlah epoch, diketahui bahwa kinerja model terbaik di seluruh percobaan yang dilakukan berhasil di dapatkan pada rasio data 50:50 dengan nilai *accuracy* sebesar 88.21% pada epoch ke-30 (lihat Tabel 5). Nilai *precision, recall* dan *f1-score* tetap konsisten di angka 88%. Meskipun pada rasio pembagian 70:30 yang tercantum pada Tabel 3 mencatat *accuracy* yang sedikit lebih tinggi, yaitu 88.23% pada epoch ke-20, tetapi rasio pembagian 50:50 dipilih sebagai yang paling optimal karena pembagian yang seimbang memberikan evaluasi yang lebih representatif. Pembagian data yang merata membantu mengurangi kemungkinan bias akibat dominasi data pelatihan dan memungkinkan model untuk diuji pada dataset uji yang setara. Selain itu, penggunaan 30 epoch pada model terbaik ini menjamin stabilitas metrik tanpa menunjukkan tanda-tanda *overfitting.* Sebaliknya, pada pembagian rasio data 30:70 (Tabel 8) menunjukkan kinerja model terendah dengan *accuracy* 85.32% pada epoch ke-10, yang mungkin disebabkan oleh jumlah data pelatihan yang lebih sedikit sehingga pola sulit untuk dipelajari oleh model.

Meskipun terdapat sedikit perbedaan, preferensi terhadap pembagian rasio data 50:50 didasarkan pada prinsip kesetaraan data, yang menjadikan evaluasi model lebih objektif dan representatif. Rasio ini memungkinkan pengujian yang menyeluruh terhadap kemampuan generalisasi model, sekaligus menghindari ketergantungan yang berlebihan pada data pelatihan.

Konsistensi dalam rentang kinerja di seluruh percobaan yang dilakukan juga menunjukkan bahwa kumpulan data yang digunakan memiliki kualitas yang baik, sehingga mendukung proses pelatihan dan evaluasi. Oleh karena itu, meskipun model menunjukkan kinerja yang relatif stabil di semua percobaan, pembagian rasio data 50:50 dipilih untuk memberikan keseimbangan yang optimal antara pelatihan dan evaluasi tanpa mengorbankan representasi data.

5. KESIMPULAN DAN SARAN

Berdasarkan temuan dari penelitian ini, dapat disimpulkan bahwa dari total 5.000 data ulasan aplikasi LinkedIn yang di analisis dengan pendekatan berbasis *Lexicon*, terdapat sebaran sentimen yang relatif seimbang. Dari jumlah tersebut, 2.755 data ulasan (55.1%) menunjukkan sentimen positif, sementara 2.245 data ulasan (44.9%) menunjukkan sentimen negatif. Temuan ini mengindikasikan bahwa kumpulan data yang digunakan mampu mencerminkan bahwa selama periode antara tahun 2020 hingga 2024, mayoritas ulasan yang diberikan untuk aplikasi LinkedIn di *Play Store* memiliki sentimen positif, meskipun selisih antara sentimen positif dan negatif tergolong kecil. Kemudian penerapan model LSTM dengan pendekatan *Machine Learning* menunjukkan hasil yang sangat baik. Dari berbagai percobaan yang telah dilakukan, kinerja model terbaik berhasil diperoleh pada rasio pembagian data 50:50, dengan *accuracy* 88.21%, *precision, recall* dan *f1-score* dengan nilai yang konsisten yaitu 88% pada epoch ke-30. Pemilihan pembagian rasio ini didasarkan pada pembagian data pelatihan dan pengujian yang seimbang, yang memungkinkan pengujian yang menyeluruh terhadap kemampuan generalisasi model, sekaligus menghindari ketergantungan yang berlebihan pada data pelatihan. Perbedaan antara berbagai rasio pembagian data ini relatif kecil, berkisar antara 85% hingga 88%.

Hal ini menunjukkan bahwa model LSTM memiliki ketahanan dan stabilitas yang bai dalam menyesuaikan diri dengan variasi jumlah data pelatihan dan pengujian. Oleh karena itu, penelitian ini menunjukkan bahwa pendekatan *Hybrid* yang diterapkan mampu memberikan kinerja analisis sentimen yang handal dan memberikan wawasan berharga mengenai persepsi pengguna terhadap aplikasi LinkedIn di *Play Store*. Di masa depan, penulis menyarankan untuk memperluas atau memperbarui sentimen leksikon yang digunakan, terutama dengan memperhatikan variasi bahasa informal yang sering digunakan dalam ulasan dan juga untuk terus mengembangkan dan menyempurnakan pemodelan LSTM, misalnya dengan penggunaan *word embedding* yang lebih canggih, atau menggabungkan model *pre-trained* agar dapat meningkatkan kemampuan dalam memahami konteks bahasa yang lebih kompleks. Serta saran untuk melakukan beberapa percobaan dalam pemodelan LSTM untuk mendapatkan kinerja model yang lebih optimal.

DAFTAR PUSTAKA

[1] S. A. Bhat and N. F. Huang, "Big Data and AI Revolution in Precision Agriculture: Survey and Challenges," *IEEE Access*, vol. 9, pp. 110209–110222, 2021, doi: 10.1109/ACCESS.2021.3102227.

[2] A. Ligthart, C. Catal, and B. Tekinerdogan, *Systematic reviews in sentiment analysis: a tertiary study*, vol. 54, no. 7. Springer Netherlands, 2021. doi: 10.1007/s10462-021-09973-3.

[3] P. Mehta and S. Pandya, "A review on sentiment analysis methodologies, practices and applications," *Int. J. Sci. Technol. Res.*, vol. 9, no. 2, pp. 601–609, 2020.

[4] 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.

[5] N. M. Sham and A. Mohamed, "Climate Change Sentiment Analysis Using Lexicon, Machine Learning and Hybrid Approaches," *Sustain.*, vol. 14, no. 8, 2022, doi: 10.3390/su14084723.

[6] S. Sharma and P. Chaudhary, "Machine learning and deep learning," *Quantum Comput. Artif. Intell. Train. Mach. Deep Learn. Algorithms Quantum Comput.*, pp. 71–84, 2023, doi: 10.1515/9783110791402-004.

[7] R. Catelli, S. Pelosi, and M. Esposito, "Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian," *Electron.*, vol. 11, no. 3, 2022, doi: 10.3390/electronics11030374.

[8] B. Agarwal and N. Mittal, "Machine learning approaches for sentiment analysis," *Artif. Intell. Concepts, Methodol. Tools, Appl.*, vol. 3, no. March, pp. 1740–1756, 2022, doi: 10.4018/978-1-5225-1759-7.ch070.

[9] Gishella Septania Al-Husna, Dian Asmarajati, Iman Ahmad Ihsannuddin, and Rina Mahmudati, "Perbandingan Metode Naïve Bayes Dan Support Vector Machine Untuk Analisis Sentimen Pada Ulasan Pengguna Aplikasi Linkedin," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 3, no. 2, pp. 139–144, 2024, doi: 10.55123/storage.v3i2.3602.

[10] M. Defriani, M. R. Muttaqin, and Q. R. Karima, "Sentiment Analysis of the LinkedIn Application Using the Lexicon Based Method Based on Google Play Store Reviews," *Inf. Syst. Technol.*, vol. 5, no. 1, pp. 1–14, 2024.

[11] S. W. Irianandа, R. W. Budiawan, A. Y. Rahman, and I. Istiadi, "Optimasi Klasifikasi Sentimen Komentar Pengguna Game Bergerak Menggunakan Svm, Grid Search Dan Kombinasi N-Gram," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 4, pp. 743–752, 2024, doi: 10.25126/jtiik.1148244.

[12] E. Damayanti, A. V. Vitianingsih, S. Kacung, H. Suhartoyo, and A. Lidya Maukar, "Sentiment Analysis of Alfagift Application User Reviews Using Long Short-Term Memory (LSTM) and Support Vector Machine (SVM) Methods," *Decod. J. Pendidik. Teknol. Inf.*, vol. 4, no. 2, pp. 509–521, 2024, doi: 10.51454/decode.v4i2.478.

[13] V. Bonta, N. Kumaresh, and N. Janardhan, "A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis," *Asian J. Comput. Sci. Technol.*, vol. 8, no. S2, pp. 1–6, 2019, doi: 10.51983/ajcst-2019.8.s2.2037.

[14] A. Chamekh, M. Mahfoudh, and G. Forestier, "Sentiment Analysis Based on Deep Learning in E-Commerce," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13369 LNAI, pp. 498–507, 2022, doi: 10.1007/978-3-031-10986-7_40.

[15] F. Koto and G. Y. Rahmaningtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," *Proc. 2017 Int. Conf. Asian Lang. Process. IALP 2017*, vol. 2018-Janua, no. December, pp. 391–394, 2017, doi: 10.1109/IALP.2017.8300625.

[16] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU," no. Ml, pp. 1–61, 2023, [Online]. Available: http://arxiv.org/abs/2305.17473

[17] D. B. Riyanto, A. Y. Rahman, and Istiadi, "Children with Speech Disorders Voice Classification: LSTM and BiLSTM Approach Based on MFCC Features," in *Proceedings: ICMERALDA 2023 - International Conference on Modeling and E-Information Research, Artificial Learning and Digital Applications*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 35–38. doi: 10.1109/ICMERALDA60125.2023.10458192

[18] R. Huang *et al.*, "A Review On Evaluation Metrics For Data Classification Evaluations," *Med. Image Anal.*, vol. 80, no. 2, p. 102478, 2022.

[19] K. T. Hanna, "What Is LinkedIn," *WhatIs.com*, 2022.

[20] M. A. Palomino, "Evaluating-the-Effectiveness-of-Text-PreProcessing-in-Sentiment-AnalysisApplied-Sciences-Switzerland.pdf," 2022.

[21] A. El Kah and I. Zeroual, "The effects of Pre-Processing Techniques on Arabic Text Classification," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 10, no. 1, pp. 41–48, 2021, doi: 10.30534/ijatcse/2021/061012021.

[22] N. Garg and K. Sharma, "Text pre-processing of multilingual for sentiment analysis based on social network data," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 1, pp. 776–784, 2022, doi: 10.11591/ijece.v12i1.pp776-784.

[23] A. Tabassum and R. R. Patil, "A Survey on Text Pre-Processing & Feature Extraction Techniques in Natural Language Processing," *Int. Res. J. Eng. Technol.*, no. June, pp. 4864–4867, 2020, [Online]. Available: www.irjet.net

[24] J. Mantik *et al.*, "Optimizing nazief adriani's stemmer algorithm in detecting indonesian word errors using sastrawi," *J. Mantik*, vol. 7, no. 3, pp. 1733–1742, 2023, [Online]. Available: https://iocscience.org/ejournal/index.php/mantik/article/view/4020

[25] M. Grohe, "Word2vec, node2vec, graph2vec, X2vec: Towards a Theory of Vector Embeddings of Structured Data," *Proc. ACM SIGACT-SIGMOD-SIGART Symp. Princ. Database Syst.*, pp. 1–16, 2020, doi: 10.1145/3375395.3387641.

[26] I. O. Muraina, "Ideal Dataset Splitting Ratios in Machine Learning Algorithms: General Concerns for Data Scientists and Data Analysts," *7th Int. Mardin Artuklu Sci. Res. Conf.*, no. February, pp. 496–504, 2022, [Online]. Available: https://www.researchgate.net/publication/358284895_IDEAL_DATASET_SPLITTING_RATIOS_IN_MACHINE_LEARNING_ALGORITH

MS_GENERAL_CONCERNS_FOR_DATA_SCIENTISTS_AND_DATA_ANALYSTS
[27] Q. H. Nguyen *et al.*, "Influence of data splitting on performance of machine learning models in prediction of shear strength of soil," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/4832864.