

## **Analisis Sentimen Publik Terhadap Enterprise Resource Planning (ERP) Di Media Sosial X Menggunakan Model Roberta Dan Twitbert**

**Berry Agustianto<sup>1</sup>, Ahmad Musyafa<sup>2</sup>, Arya Adhyaksa Waskita<sup>3</sup>**

<sup>1</sup>Magister Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang  
Jl. Surya Kencana No.1, Tangerang Selatan, Indonesia

e-mail: [beragustianto@gmail.com](mailto:beryagustianto@gmail.com)<sup>1</sup>, [dosen00668@unpam.ac.id](mailto:dosen00668@unpam.ac.id)<sup>2</sup>, [aawaskita@unpam.ac.id](mailto:aawaskita@unpam.ac.id)<sup>3</sup>

Received : July, 2025

Accepted : August, 2025

Published : August, 2025

### **Abstract**

*Enterprise Resource Planning (ERP) has become an essential part of organizational digital transformation, helping to integrate various business functions. However, public acceptance of ERP on social media often serves as a crucial indicator of its successful implementation. This study aims to analyze public sentiment toward ERP on the social media platform X using two deep learning-based natural language processing models, RoBERTa and TweetBERT. These models are used to classify sentiment into Positive, Negative, or Neutral categories, with the goal of gaining a deeper understanding of public views on ERP. In this study, data was collected from social media X with 700 data points using a crawling method to obtain a substantial number of posts related to ERP. The data was processed through several stages, including text preprocessing, tokenization, and model training. RoBERTa, known for its ability to deeply understand the context of text, was compared with TweetBERT, a model optimized for short texts such as tweets. Evaluation was conducted using accuracy, precision, recall, and F1-score metrics to assess the performance of both models. The results show that RoBERTa has an accuracy of 0.40, with the best precision for positive sentiment (0.91) and the highest recall for negative sentiment (0.76). However, RoBERTa struggles to recognize positive sentiment, with a recall of only 0.23. On the other hand, TweetBERT shows higher accuracy (0.76), with the best precision for positive sentiment (0.89) and the highest recall for neutral sentiment (0.86). Although TweetBERT is more effective in capturing positive and neutral sentiments, it struggles to identify negative sentiment (with a recall of 0.12). This study provides valuable insights for companies and stakeholders to better understand public perception of ERP and assists in planning more effective communication strategies. The research also opens opportunities for further development by using a larger dataset or combining other models to improve sentiment prediction accuracy.*

Keywords: sentiment analysis, erp, model transformer, roberta, twitbert.

### **Abstrak**

*Enterprise Resource Planning (ERP) telah menjadi bagian penting dalam transformasi digital organisasi, membantu mengintegrasikan berbagai fungsi bisnis. Namun, penerimaan publik terhadap ERP di media sosial sering kali menjadi indikator penting keberhasilan implementasinya. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap ERP di platform media sosial X menggunakan dua model pemrosesan bahasa alami berbasis deep learning, yaitu RoBERTa dan TweetBERT. Model-model ini digunakan untuk mengklasifikasikan sentimen menjadi kategori Positif, Negatif, atau Netral, dengan*

*tujuan memperoleh pemahaman mendalam mengenai pandangan publik terhadap ERP. Dalam penelitian ini, data dikumpulkan dari media sosial X sebanyak 700 data dengan menggunakan metode crawling data untuk mendapatkan sejumlah besar postingan terkait ERP. Data tersebut diproses melalui serangkaian tahap, termasuk preprocessing teks, tokenisasi, dan pelatihan model. RoBERTa, yang terkenal karena kemampuannya memahami konteks teks secara mendalam, dibandingkan dengan TweetBERT, model yang dioptimalkan untuk teks pendek seperti tweet. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score untuk mengukur kinerja kedua model. Hasil penelitian menunjukkan bahwa RoBERTa memiliki akurasi 0.40, dengan precision terbaik pada sentimen positif (0.91) dan recall tertinggi pada sentimen negatif (0.76). Namun, RoBERTa mengalami kesulitan dalam mengenali sentimen positif, dengan recall hanya 0.23. Di sisi lain, TweetBERT menunjukkan akurasi lebih tinggi (0.76), dengan precision terbaik pada sentimen positif (0.89) dan recall tertinggi pada sentimen netral (0.86). Meskipun TweetBERT lebih efektif dalam menangkap sentimen positif dan netral, model ini kesulitan mengenali sentimen negatif (recall 0.12). Penelitian ini memberikan wawasan berharga bagi perusahaan dan pemangku kepentingan untuk lebih memahami persepsi publik terhadap ERP serta membantu dalam perencanaan strategi komunikasi yang lebih efektif. Penelitian ini juga membuka peluang untuk pengembangan lebih lanjut dengan menggunakan dataset yang lebih luas atau menggabungkan model-model lain untuk meningkatkan akurasi prediksi sentimen.*

**Kata Kunci:** analisis sentimen, ERP, model transformer, roberta, twitbert.

## 1. PENDAHULUAN

Enterprise Resource Planning (ERP) adalah sistem yang dirancang untuk menyatukan berbagai fungsi bisnis, seperti keuangan, logistik, dan manajemen sumber daya manusia, dalam satu platform tunggal terpadu [1]. Implementasi ERP dapat meningkatkan efisiensi operasional, menghemat biaya, dan mempercepat pengambilan keputusan berbasis data [2]. Namun, keberhasilan penerapannya tidak hanya bergantung pada aspek fungsionalitas teknis, tetapi juga pada sejauh mana penerimaan dan pandangan pengguna terhadap sistem tersebut. Di era digital sekarang, media sosial telah menjadi platform utama bagi pengguna untuk berbagi pandangan dan pengalaman mereka tentang teknologi seperti ERP. Oleh karena itu, Analisis Sentimen terhadap ERP di media sosial dapat menjadi indikator penting dalam mengevaluasi keberhasilan implementasi sistem ini [3].

Penelitian sebelumnya menunjukkan bahwa opini pengguna di platform digital memiliki pengaruh besar terhadap adopsi teknologi [4]. Menemukan bahwa mayoritas pengguna yang berbagi pengalaman negatif tentang ERP di media sosial dapat mempengaruhi persepsi organisasi lain yang sedang mempertimbangkan teknologi ini. Penelitian lain oleh [5] mengindikasikan bahwa analisis sentimen berbasis media sosial mampu memberikan wawasan berharga bagi pengambil

keputusan, meskipun ada tantangan dalam menangkap konteks bahasa tidak terstruktur yang sering ditemukan pada platform seperti media sosial.

Model berbasis Deep Learning seperti RoBERTa dan TweetBERT telah terbukti efektif dalam memahami teks kompleks. RoBERTa unggul dalam menangkap konteks bahasa yang dalam dan panjang, sementara TweetBERT lebih baik dalam menganalisis teks pendek dan tidak baku seperti tweet. Penelitian oleh [6] dalam analisis sentimen, perbandingan antara keduanya menunjukkan bahwa RoBERTa memiliki akurasi yang lebih baik dalam mendeteksi sentimen netral, sementara TweetBERT lebih unggul dalam mengidentifikasi sentimen positif dan negatif.

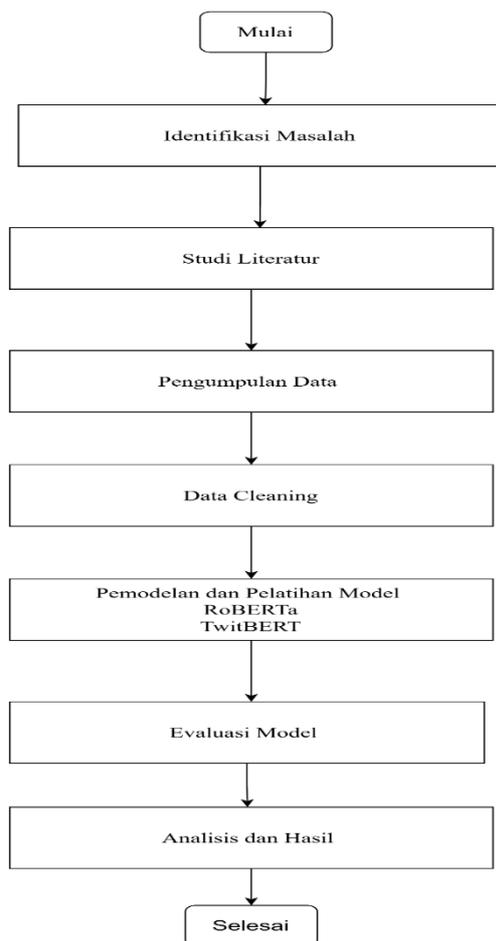
Berdasarkan temuan ini, Penelitian ini bertujuan untuk mengevaluasi sentimen masyarakat terhadap ERP di platform media sosial X kombinasi model RoBERTa dan TweetBERT. Dengan memanfaatkan keunggulan masing-masing model, diharapkan, penelitian ini dapat menghasilkan temuan yang lebih tepat dan mendalam mengenai pandangan publik terhadap ERP. Penelitian ini juga berkontribusi pada pengembangan metode analisis sentimen berbasis NLP dan memberikan wawasan praktis bagi perusahaan untuk menyusun strategi

komunikasi yang lebih efektif terkait implementasi ERP.

## 2. METODE

### 2.1 Kerangka Pemikiran

Pada gambar 2.1 menggambarkan metode penelitian yang digunakan untuk melakukan analisis sentimen.



Gambar 2. 1 Kerangka Pemikiran

#### 2.1.1 Identifikasi Masalah

Penelitian ini berfokus pada kurangnya pemahaman persepsi publik mengenai sistem ERP di media sosial, dengan analisis sentimen yang akurat dan relevan. Di era digital, sistem Enterprise Resource Planning memegang peranan krusial dalam meningkatkan efisiensi dan produktivitas bisnis [7]. Namun, efektivitas sistem ERP sangat bergantung pada implementasi dan pemanfaatannya yang tepat, karena implementasi yang tidak tepat dapat menimbulkan gangguan terhadap proses bisnis, terutama bagi Usaha Mikro, Kecil, dan Menengah yang memiliki keterbatasan sumber daya manusia

dan finansial [8]. Implementasi sistem informasi dalam proses bisnis perusahaan dapat meningkatkan efisiensi dan efektivitas, tetapi tanpa perencanaan yang baik, implementasi tersebut mungkin tidak memberikan dampak yang diharapkan [9].

#### 2.1.2 Studi Literatur

Kajian dilakukan tentang analisis sentimen, model berbasis Transformer seperti RoBERTa dan TwitBERT, serta penerapan sistem ERP, untuk membangun dasar teoretis dan metodologis.

#### 2.1.3 Pengumpulan Data

Data dikumpulkan dari media sosial X (Twitter) dengan kata kunci "ERP," menghasilkan tweet yang mencerminkan opini publik tentang sistem ERP.

#### 2.1.4 Data Cleaning

Data dibersihkan dari elemen tidak relevan (URL, mention, tanda baca, karakter khusus) agar siap diproses oleh model analisis sentimen.

#### 2.1.5 Pemodelan

Model RoBERTa dan TwitBERT yang telah di-fine-tune diterapkan untuk klasifikasi sentimen (Positif, Netral, Negatif), dengan pembandingan menggunakan pendekatan leksikon VADER.

#### 2.1.6 Evaluasi Model

Kinerja model dievaluasi dengan metrik seperti akurasi, precision, recall, F1-score, serta confusion matrix dan classification report.

#### 2.1.7 Analisis Dan Hasil

Hasil analisis menunjukkan distribusi sentimen dan perbandingan akurasi antar model, yang digunakan untuk menyimpulkan model yang paling tepat untuk menangkap sentimen publik terkait ERP di media sosial.

### 2.2 Pengumpulan Data

Dalam penelitian ini, data dikumpulkan dengan menggunakan pendekatan web crawling melalui Python, yang memungkinkan akses dan ekstraksi konten secara otomatis dari platform media sosial X (sebelumnya Twitter). Crawler mengumpulkan data berdasarkan kata kunci "Review ERP lang:id" untuk menyaring konten berbahasa Indonesia yang relevan dengan pengetahuan masyarakat

mengenai ERP. Data yang dikumpulkan berupa unggahan dan komentar pengguna yang kemudian dianalisis menggunakan model transformer, yakni RoBERTa dan TweetBERT. Kedua model ini digunakan untuk melakukan analisis sentimen terhadap data teks, memberikan klasifikasi yang lebih akurat dan mendalam. Proses ini mendukung evaluasi berbasis data dalam konteks implementasi ERP.

Gambar 2. 2 Sample Data

Pada Gambar 1, data yang berhasil diperoleh dari platform media sosial X berjumlah 700 entri dan diekspor ke dalam format CSV, mencakup ulasan dan tanggapan pengguna terkait implementasi ERP. Selanjutnya, data tersebut akan melalui tahapan preprocessing untuk membersihkan dan menyiapkan teks agar siap dianalisis sentimennya.

### 2.3 Preprocessing

Proses preprocessing mencakup beberapa tahapan penting, seperti normalisasi huruf (casing normalization), tokenisasi untuk memecah kalimat menjadi unit kata, pembersihan karakter non-alfabetik (seperti simbol dan tanda baca), dan penghapusan stopwords [10]. Tahapan preprocessing data merupakan langkah penting yang dilakukan setelah pengumpulan data [11], penghapusan stopwords atau Kata-kata umum yang tidak memberikan informasi signifikan (seperti "yang", "dan", "di"), serta penerapan stemming atau lemmatisasi untuk

Tabel 2.1 Hasil Pelabelan Sentimen

Cleaned Text	Actual sentiment	Sentiment RoBERTa	Sentiment twitbert
ERP system is well equipped with modern technology and artificial intelligence system which covers all aspects of modern world requirements contact us today infocom erpsystem ERP isolat ERP crm customerrelationshipmanagement	POSITIVE	NEGATIVE	POSITIVE

mengembalikan kata pada bentuk dasarnya. Langkah-langkah ini penting untuk meningkatkan kualitas data agar dapat diproses secara maksimal oleh model analisis sentimen.

### 2.4 Distribusi Sentimen

distribusi sentimen berdasarkan pelabelan awal menggunakan pendekatan leksikon VADER, Yang

mengelompokkan data ke dalam tiga kategori: Positif, Negatif, dan Netral [12]. Selanjutnya, pada tahap preprocessing, berbagai teknik diterapkan untuk mempersiapkan data teks sebelum analisis lebih lanjut [11]. Transformasi ini sangat penting untuk memastikan kualitas dan konsistensi data, yang pada gilirannya memengaruhi kinerja model analisis sentimen secara keseluruhan [13]. Dari total 700 tweet yang dianalisis, mayoritas termasuk dalam kategori sentimen Positif, dengan jumlah 445 tweet (62,41%). Sentimen Netral tercatat sebanyak 243 tweet (34,08%), sementara hanya 25 tweet (3,51%) yang dikategorikan sebagai sentimen Negatif yang ditunjukkan pada gambar 2.3.

Gambar 2. 3 Hasil Distribusi Sentimen VADER

Tabel 2.1 berikut menyajikan beberapa contoh tweet yang dianalisis dalam penelitian ini beserta hasil pelabelan sentimennya, baik berdasarkan metode leksikon (VADER) yang digunakan sebagai Label Aktual dan sudah di tambahkan kata-kata indonesia baru ke lexicon[14], maupun hasil prediksi dari model RoBERTa dan TwitBERT.

tired of tedious data management streamline and simplify with synthe grate solutions ERP the smarter way to centralize your schools data learn more contact us 923330792076 schooldatamanagement erpsystem edtechsolutions smartschools	POSITIVE	NEGATIVE	POSITIVE
what s the most important factor when choosing an ERP system 91 8369491623 bangladesh syria growthawk businessdevelopment softwaredevelopment erpsoftware ERP polls polloftheday birth rc16 cyclonefengal questiontime erpsystem	NEGATIVE	NEGATIVE	NEUTRAL
what s the most important factor when choosing an ERP system 91 8369491623 bangladesh syria growthawk businessdevelopment softwaredevelopment erpsoftware ERP polls polloftheday birth rc16 cyclonefengal questiontime erpsystem	POSITIVE	POSITIVE	NEUTRAL

## 2.5 Model Algoritma

### 2.5.1 RoBERTa

*RoBERTa* adalah alternatif/penerus *BERT* yang ditingkatkan dengan mengoptimalkan hyperparameter pelatihan untuk *BERT* dan membuatnya mengungguli *BERT* pada hampir semua tugas yang dirancang untuk diselesaikan oleh *BERT* [15][16]. *RoBERTa* secara harfiah menggunakan arsitektur yang sama dengan *BERT*. Tidak seperti *BERT*, selama pre-training *RoBERTa* hanya dilatih dengan *Masked Language Modeling (BERT* juga dilatih dengan *Masked Language Modelling* dan *Next Sentence Prediction*).

### 2.5.2 TwitBERT

*BERT* (2018) merupakan model berbasis *Transformer* yang diperkenalkan oleh Google. *BERT* merevolusi *NLP* dengan kemampuannya memahami konteks dua arah dalam teks, menjadikannya landasan untuk model lain, termasuk *TwitBERT* [17]. Setelah *BERT* menjadi populer, para peneliti mulai membuat variasi domain-spesifik, misalnya *BioBERT* untuk biomedis dan *FinBERT* untuk keuangan. *TwitBERT* muncul sebagai respons terhadap kebutuhan analisis data dari media sosial. Data *Twitter* yang kaya dengan teks informal dan kontekstual memerlukan adaptasi khusus dari model *BERT*.

*TwitBERT* adalah model pemrosesan bahasa alami (*Natural Language Processing, NLP*) berbasis *Transformer* yang dirancang khusus untuk memahami dan menganalisis teks dari platform media sosial, terutama *Twitter*. Model ini memanfaatkan arsitektur *Bidirectional Encoder Representations from Transformers*, yang terkenal karena kemampuannya menangkap konteks dan nuansa linguistik dalam teks [18]. Kemampuan untuk mempertimbangkan kata-kata sebelum dan sesudah kata target memungkinkan *TwitBERT* memahami teks dengan lebih mendalam, sangat penting mengingat karakteristik unik komunikasi media sosial, termasuk sifat informal, kelimpahan jargon, dan kecenderungan untuk menggunakan bahasa yang ambigu. Sebagai model yang dirancang untuk data tekstual media sosial, *TwitBERT* menjanjikan peningkatan akurasi dan pemahaman yang lebih mendalam tentang data tekstual *Twitter*, yang membuka prospek baru untuk penelitian dan aplikasi yang menjangkau berbagai disiplin ilmu [19]. Media sosial, khususnya *Twitter*, telah menjadi pusat penyebaran informasi dan opini publik yang dinamis, sehingga membutuhkan alat yang canggih untuk analisis data tekstual [20].

Model ini merupakan varian dari *BERT (Bidirectional Encoder Representations from*

Transformers) yang disesuaikan untuk menangani ciri khas teks *Twitter*, seperti singkatan, emoji, hashtag, dan panjang teks yang terbatas.

## 2.6 Evaluasi Hasil

Berbagai teknik evaluasi yang umum digunakan untuk menilai kinerja model klasifikasi meliputi Precision, Recall, dan F1-Score. Tujuan dari pengukuran ini adalah untuk menghasilkan model dengan kesalahan klasifikasi seminimal mungkin serta meminimalkan ketidakpastian dalam hasil prediksi [21]. Setelah model diuji menggunakan data uji, langkah selanjutnya adalah proses evaluasi yang bertujuan untuk menilai sejauh mana model dapat memprediksi kelas dengan tepat dan akurat. Proses evaluasi ini mencakup perhitungan nilai Precision untuk melihat rasio prediksi positif yang benar, Recall untuk mengukur sejauh mana model dapat mengidentifikasi semua kasus positif, dan F1-Score sebagai rata-rata harmonis antara Precision dan Recall, yang memberikan gambaran menyeluruh tentang keseimbangan kinerja model dalam mendeteksi kelas target [22].

## 3. HASIL DAN PEMBAHASAN

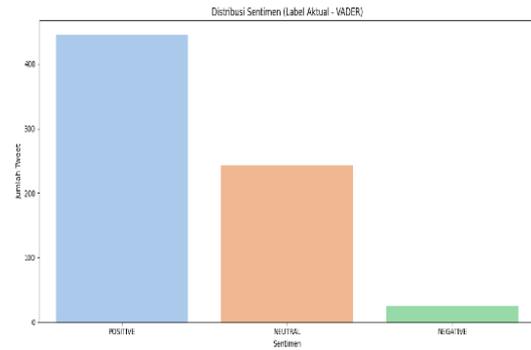
Dataset dalam penelitian ini menggunakan jenis format CSV yang dikumpulkan melalui metode Web crawling dilakukan pada media sosial X. Data yang dikumpulkan berupa tweet dengan menggunakan kata kunci "Review ERP lang:id" Pengambilan data dilakukan secara acak, mencakup tweet dari pengguna biasa maupun Platform media online di X. Jumlah total dataset yang digunakan adalah 700 tweet pengumpulan data dalam jumlah besar dari platform media sosial X memerlukan akses API yang terbatas oleh karena itu jumlah tweet yang dapat dikumpulkan penelitian ini dibatasi oleh faktor waktu sehingga penelitian ini bersifat eksploratif dan berfungsi sebagai dasar untuk penelitian lebih lanjut. Untuk pengolahan data nya menggunakan bahasa pemrograman python.

### 3.1 Hasil Pelabelan Sentimen

#### 3.1.1 Hasil Klasifikasi dengan VADER

Hasil klasifikasi sentimen disajikan menggunakan pendekatan berbasis leksikon sebagai pembandingan dasar (*actual sentiment*). Analisis awal dilakukan dengan VADER untuk memberikan label sentimen (Label Aktual), yang mengevaluasi polaritas setiap tweet dan mengkategorikannya ke dalam tiga kelas: *Positif*, *Netral*, dan *Negatif*.

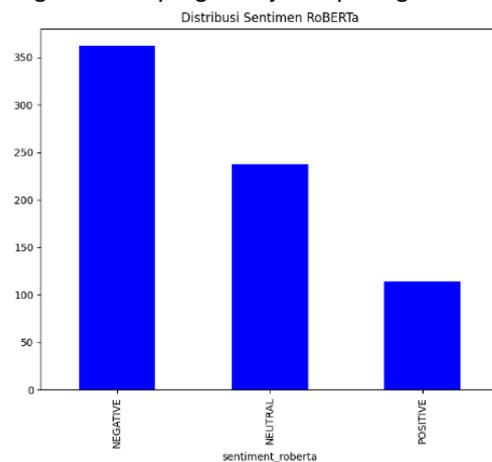
Sentimen positif mendominasi dengan 445 tweet, diikuti oleh netral 243 tweet dan sentimen negatif 25 tweet yang ditunjukkan pada Gambar 3.1.



Gambar 3. 1 Hasil Distribusi Sentimen VADER

#### 3.1.2 Hasil Klasifikasi dengan RoBERTa

Selanjutnya, tweet yang telah dibersihkan dianalisis menggunakan model RoBERTa. Model ini melakukan klasifikasi berdasarkan emosi yang kemudian dimapping ke label sentimen: joy, love, surprise → Positif, anger, fear, sadness → Negatif, dan sisanya dianggap Netral. Sebagian besar tweet diklasifikasikan sebagai Negatif sebanyak 362 tweet, diikuti oleh Netral sebanyak 237 tweet, dan hanya 114 tweet yang termasuk kategori Positif yang ditunjukkan pada gambar 3.2.

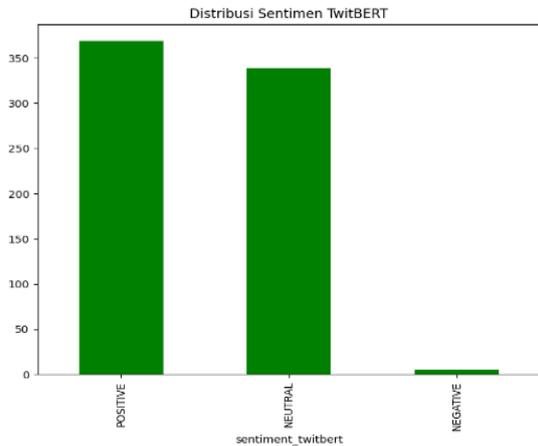


Gambar 3. 2 Distribusi Sentimen RoBERTa

#### 3.1.3 Hasil Klasifikasi dengan TwitBERT

Model yang digunakan selanjutnya adalah TwitBERT, yang dioptimalkan untuk data Twitter dan mengklasifikasikan sentimen menjadi tiga kategori: POS, NEU, dan NEG, yang kemudian dipetakan ke POSITIVE, NEUTRAL, dan NEGATIVE. TwitBERT, yang dilatih khusus untuk media sosial, menunjukkan distribusi prediksi yang berbeda

signifikan dari RoBERTa, dengan 369 tweet diprediksi Positif, 339 tweet Netral, dan 5 tweet Negatif. Pada gambar 3.3 Pola ini menunjukkan bahwa TwitBERT cenderung memberi label Positif pada sebagian besar data, sesuai dengan karakteristik tweet yang ekspresif dan tidak formal.



Gambar 3. 3 Distribusi Sentimen TwitBERT

### 3.2 Evaluasi Performa Model

#### 3.2.1 Confusion Matrix

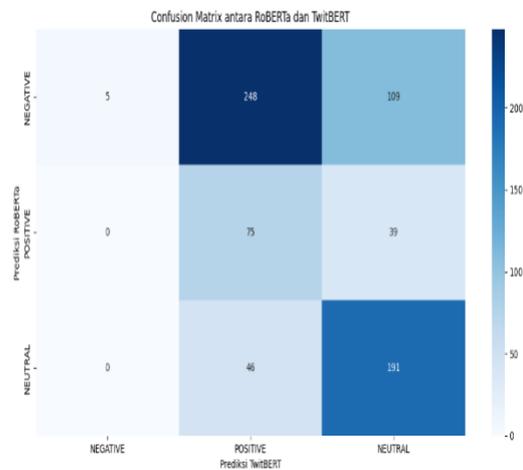
Confusion matrix berikut membandingkan hasil klasifikasi sentimen antara dua model berbasis transformer, RoBERTa dan TwitBERT. Pada matrix ini, prediksi RoBERTa digunakan sebagai acuan di sumbu vertikal (baris), sedangkan prediksi TwitBERT berada di sumbu horizontal (kolom). Diagonal matrix menunjukkan kesesuaian prediksi antara kedua model:

1. 5 tweet dengan sentimen Negatif
2. 75 tweet dengan sentimen Positif
3. 191 tweet dengan sentimen Netral

Namun, terdapat perbedaan signifikan, terutama pada kelas Negatif. RoBERTa lebih cenderung mengklasifikasikan tweet sebagai Negatif, dengan 248 tweet yang oleh TwitBERT dikategorikan sebagai Positif dan 109 tweet sebagai Netral. Hal ini mengindikasikan bahwa RoBERTa memiliki sensitivitas lebih tinggi terhadap sentimen Negatif.

Selain itu, terdapat ketidaksesuaian pada kelas Netral, di mana 46 tweet yang diklasifikasikan Netral oleh RoBERTa justru dianggap Positif oleh TwitBERT. Perbedaan ini mencerminkan karakteristik pelatihan masing-masing model: RoBERTa lebih konservatif dalam mengklasifikasikan sentimen Positif, sementara

TwitBERT, yang dilatih khusus untuk teks Twitter, lebih adaptif terhadap ekspresi informal. Secara keseluruhan, meski ada kesepahaman antara kedua model, perbedaan dalam klasifikasi sentimen, khususnya antara Negatif dan Positif, perlu diperhatikan yang ditampilkan pada Gambar 3.4.



Gambar 3. 4 Confusion Matrix RoBERTa dan TwitBERT

### 3.3 Classification Report

#### 3.3.1 RoBERTa

Pada Gambar 3.5 Model RoBERTa menghasilkan akurasi 40,39%, yang mencerminkan performa klasifikasi yang kurang optimal. Pada kategori Positif, precision tinggi (0.91), namun recall rendah (0.23), menunjukkan model cukup yakin dalam memprediksi Positif, namun hanya sedikit tweet Positif yang terdeteksi, dengan f1-score 0.37. Untuk kategori Netral, precision (0.70) dan recall (0.68) cukup seimbang, dengan f1-score 0.69. Pada kategori Negatif, recall tinggi (0.76), namun precision sangat rendah (0.05), menghasilkan f1-score rendah 0.10. Nilai rata-rata makro f1-score 0.39 menunjukkan performa model yang belum merata di seluruh kelas.

```

Classification Report RoBERTa:
precision    recall  f1-score   support

NEGATIVE     0.05    0.76    0.10        25
NEUTRAL      0.70    0.68    0.69       243
POSITIVE     0.91    0.23    0.37       445

accuracy          0.40        713
macro avg         0.55    0.56    0.39        713
weighted avg      0.81    0.40    0.47        713
    
```

Gambar 3. 5 Classification Report RoBERTa

#### 3.3.2 TwitBERT

Berbeda dengan RoBERTa, model TwitBERT menunjukkan kinerja yang lebih unggul dengan

akurasi sebesar 76,02%. Untuk kategori Positif, precision 0,89, recall 0,74, dan f1-score 0,81 mengindikasikan bahwa model ini memiliki kemampuan yang handal dalam mendeteksi sentimen Positif. Untuk kategori Netral, f1-score tinggi (0.72), recall baik (0.86), meskipun precision lebih rendah (0.62), menunjukkan efektivitas TwitBERT dalam mengidentifikasi tweet Netral. Pada kategori Negatif, precision 0.60 dan recall 0.12 menghasilkan f1-score rendah (0.20), menunjukkan kesulitan model dalam mengenali tweet Negatif. Nilai macro average f1-score 0.58 dan weighted average f1-score 0.76 menunjukkan performa yang stabil dan seimbang, terutama pada kelas Positif dan Netral yang ditunjukkan pada gambar 3.6.

```

Classification Report TwitBERT:
              precision    recall  f1-score   support

  NEGATIVE      0.60      0.12      0.20      25
   NEUTRAL      0.62      0.86      0.72     243
   POSITIVE      0.89      0.74      0.81     445

 accuracy              0.76      713
 macro avg           0.70      0.57      0.58      713
 weighted avg        0.79      0.76      0.76      713
  
```

Gambar 3. 6 Classification Report TwitBERT

### 3.4 Tabel Perbandingan Hasil Classification Report

Pada tabel 3.1 ditampilkan hasil perbandingan classification report.

Tabel 3.1 Tabel Perbandingan Hasil Classification Report

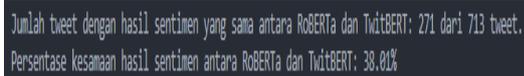
Classification	Matric	Roberta	Twitbert
Negative	precision	0,05	0,6
	Recall	0,76	0,12
	f1-score	0,1	0,2
	support	25	25
Neutral	precision	0,7	0,62
	Recall	0,68	0,86
	f1-score	0,69	0,72
	support	243	243
Positive	precision	0,91	0,89
	Recall	0,23	0,74
	f1-score	0,37	0,81
	support	445	445
Accuracy	f1-score	0,4	0,76
	support	713	713
Macro avg	precision	0,55	0,7
	Recall	0,56	0,57
	f1-score	0,39	0,58
	support	713	713
Weighted avg	precision	0,81	0,79
	Recall	0,4	0,76
	f1-score	0,47	0,76
	support	713	713

### 3.5 Perbandingan Akurasi Model

Pada gambar 3.7 Analisis kesamaan hasil prediksi antara model RoBERTa dan TwitBERT menunjukkan

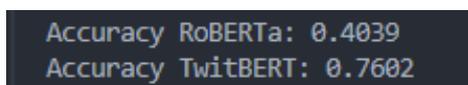
bahwa dari 713 tweet yang dianalisis, hanya 271 tweet (38,01%) yang memiliki label sentimen sama. Angka ini mencerminkan tingkat konsistensi yang

rendah, mengindikasikan perbedaan pendekatan dalam mengklasifikasikan sentimen meskipun kedua model berbasis transformer. Perbedaan ini mungkin disebabkan oleh arsitektur dan data pelatihan masing-masing model. RoBERTa adalah model general-purpose yang dilatih pada berbagai korpus teks, sementara TweetBERT dioptimalkan untuk data media sosial, membuatnya lebih cocok untuk karakteristik bahasa informal dan singkat khas Twitter.



Gambar 3. 7 Perbandingan Akurasi Model

Hasil evaluasi menunjukkan bahwa TweetBERT memiliki akurasi lebih tinggi (76,02%) dibandingkan RoBERTa (40,39%) dalam klasifikasi sentimen tweet yang ditunjukkan pada gambar 3.8. TweetBERT lebih efektif karena dirancang khusus untuk memahami bahasa informal Twitter, sedangkan RoBERTa, yang dilatih pada teks formal, kurang optimal untuk data tweet. Dengan demikian, TweetBERT lebih superior dan lebih cocok diterapkan untuk analisis sentimen di media sosial.



Gambar 3.8 Akurasi RoBERTa dan TweetBERT

#### 4. Kesimpulan

Penelitian ini bertujuan untuk mengevaluasi sentimen masyarakat terhadap Enterprise Resource Planning (ERP) di platform media sosial X dengan memanfaatkan dua model pemrosesan bahasa alami yang berbasis deep learning, yaitu RoBERTa dan TweetBERT. Hasil penelitian menunjukkan bahwa sentimen publik terhadap ERP mayoritas berada pada kategori netral, diikuti dengan sentimen positif dan negatif. Persepsi publik ini menjadi indikator penting dalam menilai keberhasilan implementasi ERP oleh perusahaan. Dari perbandingan kinerja kedua model, RoBERTa memiliki akurasi lebih rendah (0.40), namun memiliki presisi yang lebih baik pada sentimen positif dan recall tertinggi pada sentimen negatif. Di sisi lain, TweetBERT menunjukkan akurasi lebih tinggi (0.76) dengan performa terbaik pada sentimen positif dan netral, Namun, mengalami kesulitan dalam mengidentifikasi sentimen negatif. Hal ini mengindikasikan bahwa setiap model

memiliki keunggulan dalam konteks yang berbeda, di mana RoBERTa lebih efektif dalam menangani sentimen negatif dan konteks yang lebih mendalam, sementara TweetBERT lebih unggul dalam menangani teks pendek dan tidak terstruktur. Bagi perusahaan, temuan ini dapat digunakan untuk merancang strategi komunikasi yang lebih terfokus pada penguatan persepsi positif terhadap ERP. Mengingat sebagian besar opini publik sudah positif, perusahaan dapat lebih menekankan manfaat ERP melalui kanal media sosial, memanfaatkan opini positif untuk mendorong adopsi yang lebih luas. Selain itu, perusahaan perlu menangani sentimen negatif dengan pendekatan yang lebih terbuka dan responsif, mengingat adanya sebagian kecil sentimen negatif. Pada penelitian selanjutnya diharapkan dapat diperbanyak untuk jumlah data yang di crawling.

#### DAFTAR PUSTAKA

- [1] M. Meiryani, E. Fernando, S. P. Hendratno, K. Kriswanto, and S. Wifasari, 'Enterprise Resource Planning Systems: The Business Backbone', pp. 43–48, 2021, doi: 10.1145/3466029.3466049.
- [2] Md. S. A. Chowdhury, M. T. Rahman, A. M. Shahabuddin, M. R. Hassan, and M. S. R. Chowdhury, 'Implementation of Enterprise Resource Planning (ERP) in Bangladesh -Opportunities and Challenges', *International Journal of Business and Management*, vol. 16, no. 11, p. 1, 2021, doi: 10.5539/ijbm.v16n11p1.
- [3] S. AboAbdo, A. Aldhoiena, and H. Al-Amrib, 'Implementing Enterprise Resource Planning ERP System in a Large Construction Company in KSA', *Procedia Comput Sci*, vol. 164, pp. 463–470, 2019, doi: 10.1016/j.procs.2019.12.207.
- [4] M. Rizki, S. Basuki, and Y. Azhar, 'Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory(LSTM) Untuk Prediksi

- Curah Hujan Kota Malang', 2020, *Universitas Muhammadiyah Malang*. doi: 10.22219/repositor.v2i3.470.
- [5] S. Syafrudin, R. A. Nugraha, K. Handayani, S. Linawati, and W. Gata, 'Prediksi Status Pinjaman Bank dengan Deep Learning Neural Network', 2021, *Universitas Bina Sarana Informatika*. doi: 10.31294/jtk.v7i2.10474.
- [6] F. D. Pratama and H. D. Bhakti, 'IMPLEMENTASI APLIKASI PREDIKSI KETEPATAN PEMBAYARAN CUSTOMER PERUSAHAAN DENGAN METODE DECISION TREE', 2023, *Universitas Muhammadiyah Gresik*. doi: 10.30587/indexia.v5i01.5082.
- [7] R. I. Tampubolon, 'Business Improvement Strategy for Local ERP Companies in Indonesia: SWOT Method, Fuzzy AHP-TOPSIS', *Jurnal Penelitian Medan Agama*, vol. 15, no. 1, p. 84, 2024, doi: 10.58836/jpma.v15i1.21173.
- [8] W. R. Faranita and M. Si. Ir. E. Nugroho, 'IMPLEMENTATION OF THE ENTERPRISE RESOURCE PLANNING (ERP) SYSTEM ON MICRO, SMALL AND MEDIUM ENTERPRISES (MSME) BUSINESS ACTORS', *JBTI Jurnal Bisnis Teori dan Implementasi*, vol. 12, no. 2, pp. 86–93, 2021, doi: 10.18196/jbti.v12i2.12191.
- [9] D. E. Prasetyo and A. F. Wijaya, 'Information System Strategic Planning For Tourism Transportation Company Using Ward And Peppard Methodology', *INTENSIF Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 5, no. 1, pp. 43–57, 2021, doi: 10.29407/intensif.v5i1.14609.
- [10] G. P. Aulia, T. Widiharih, and I. T. Utami, 'PENERAPAN TEXT MINING DAN FUZZY C-MEANS CLUSTERING UNTUK IDENTIFIKASI KELUHAN UTAMA PELANGGAN PDAM TIRTA MOEDAL KOTA SEMARANG', *Jurnal Gaussian*, vol. 12, no. 1, pp. 126–135, 2022, doi: 10.14710/j.gauss.12.1.126-135.
- [11] I. Salamah and S. Suroso, 'Implementasi Algoritma Naive Bayes Terhadap Klasifikasi Jenis Pertanyaan Pada Perancangan Chatbot Untuk Aplikasi Penjualan Songket', *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 3, p. 1734, 2024, doi: 10.30865/mib.v8i3.7908.
- [12] L. A. Palinkas, S. M. Horwitz, C. A. Green, J. P. Wisdom, N. Duan, and K. Hoagwood, 'Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research', 2013, *Springer Science+Business Media*. doi: 10.1007/s10488-013-0528-y.
- [13] M. Zubair, J. Ali, M. Alhussein, S. Hassan, K. Aurangzeb, and M. Umair, 'An Improved Machine Learning-Driven Framework for Cryptocurrencies Price Prediction With Sentimental Cautioning', *IEEE Access*, vol. 12, pp. 51395–51418, 2024, doi: 10.1109/access.2024.3367129.
- [14] Y. Asri, W. N. Suliyanti, D. Kuswardani, and M. Fajri, 'Pelabelan Otomatis Lexicon Vader dan Klasifikasi Asri, Y., Suliyanti, W. N., Kuswardani, D., & Fajri, M. (2022). Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam menganalisis sentimen data ulasan PLN Mobile. PETIR, 15(2), 264–275. https://', *Petir*, vol. 15, no. 2, pp. 264–275, 2022.
- [15] E. Lin, J. Sun, H. Chen, and M. H. Mahoor, 'Data Quality Matters: Suicide Intention Detection on Social Media Posts Using a RoBERTa-CNN Model', *arXiv (Cornell University)*, 2024, doi: 10.48550/arxiv.2402.02262.

- [16] M. Muffo and E. Bertino, 'BERTino: an Italian DistilBERT model', in *Accademia University Press eBooks*, Accademia University Press, 2020, pp. 317–322. doi: 10.4000/books.aaccademia.8748.
- [17] J. Devlin, M. Chang, K. Lee, and K. Toutanova, 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding', *arXiv (Cornell University)*, 2018, doi: 10.48550/arXiv.1810.04805.
- [18] A. F. Hidayatullah, R. A. A. H. M. Apong, D. T. C. Lai, and A. Qazi, 'Corpus creation and language identification for code-mixed Indonesian-Javanese-English Tweets', *PeerJ Comput Sci*, vol. 9, 2023, doi: 10.7717/peerj-cs.1312.
- [19] T. Wu, Y. Wang, and N. Quach, 'Advancements in Natural Language Processing: Exploring Transformer-Based Architectures for Text Understanding', 2025, doi: 10.48550/ARXIV.2503.20227.
- [20] S. Shevira, I. M. A. D. Suarjaya, and P. W. Buana, 'Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia', *JITTER Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, no. 2, p. 1074, 2022, doi: 10.24843/jtrti.2022.v03.i02.p06.
- [21] H. GhorbanTanhaei, P. Boozary, S. Sheykhan, M. Rabiee, F. Rahmani, and I. Hosseini, 'Predictive Analytics in Customer Behavior: Anticipating Trends and Preferences', *Results in Control and Optimization*, p. 100462, 2024, doi: 10.1016/j.rico.2024.100462.
- [22] R. S. Samir, 'EfficientNet Algorithm for Classification of Different Types of Cancer', *arXiv (Cornell University)*, 2023, doi: 10.48550/arXiv.2304.08715.