

---

**KOMPARASI KINERJA MODEL SUPPORT VECTOR MACHINE DAN  
INDONESIAN BIDIRECTIONAL ENCODER REPRESENTATIONS  
FROM TRANSFORMERS DALAM KLASIFIKASI SENTIMEN  
ULASAN APLIKASI WONDR BY BNI**

**Rina Rahma Cahya<sup>1</sup>, Arief Ichwani<sup>2</sup>**

**Universitas Esa Unggul Jakarta**

e-mail: [rinarahmacahya15012@esaunggul.ac.id](mailto:rinarahmacahya15012@esaunggul.ac.id)<sup>1</sup>, [arief.ichwani@esaunggul.ac.id](mailto:arief.ichwani@esaunggul.ac.id)<sup>2</sup>

**Abstract:** *User reviews of the Wondr By BNI application on the Google Play Store contain sentiment information that can serve as a basis for evaluating the quality of digital banking services. This study aims to analyze and compare the performance of the Support Vector Machine (SVM) model and Indonesian Bidirectional Encoder Representations from Transformers (IndoBERT) in classifying user review sentiments into three categories: positive, neutral, and negative. Data was collected through web scraping techniques from the Google Play Store during the period of January–December 2025, then processed through preprocessing stages and labeled using the Lexicon-Based method with the InSet dictionary. The SVM model uses TF-IDF feature extraction with variations in the C parameter, while IndoBERT is optimized through a fine-tuning process. Evaluation is carried out using accuracy, precision, recall, and F1-score metrics. Test results show that the SVM model with an optimal parameter of C = 10 achieves an accuracy of 83.29%, whereas IndoBERT reaches an accuracy of 92.70%. IndoBERT outperforms SVM across all evaluation metrics with an average difference of around 9–10%, indicating IndoBERT's ability to understand the context of the Indonesian language more deeply. In conclusion, IndoBERT is a more effective model and is recommended for sentiment classification of Indonesian language digital banking application reviews.*

**Keywords:** *Sentiment Analysis, Support Vector Machine, IndoBERT, BERT, Wondr By BNI, Text Classification*

**Abstrak:** Ulasan pengguna aplikasi Wondr By BNI di Google Play Store mengandung informasi sentimen yang dapat menjadi bahan evaluasi kualitas layanan perbankan digital. Penelitian ini bertujuan menganalisis dan membandingkan kinerja model Support Vector Machine (SVM) dan Indonesian Bidirectional Encoder Representations from Transformers (IndoBERT) dalam mengklasifikasikan sentimen ulasan pengguna ke dalam tiga kategori: positif, netral, dan negatif. Data dikumpulkan melalui teknik web scraping dari Google Play Store pada periode Januari–Desember 2025, kemudian diproses melalui tahapan preprocessing dan dilabeli menggunakan metode Lexicon-Based dengan kamus InSet. Model SVM menggunakan ekstraksi fitur TF-IDF dengan variasi nilai parameter C, sedangkan IndoBERT dioptimalkan melalui proses fine-tuning. Evaluasi dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model SVM dengan parameter optimal C = 10 memperoleh akurasi sebesar 83,29%, sedangkan IndoBERT mencapai akurasi sebesar 92,70%. IndoBERT mengungguli SVM pada seluruh metrik evaluasi dengan selisih rata-rata sekitar 9–10%, yang menunjukkan kemampuan IndoBERT dalam memahami konteks bahasa Indonesia secara lebih mendalam. Kesimpulannya, IndoBERT merupakan model yang lebih efektif dan direkomendasikan untuk klasifikasi sentimen ulasan aplikasi perbankan digital berbahasa Indonesia.

**Kata Kunci:** Analisis Sentimen, Support Vector Machine, IndoBERT, BERT, Wondr By BNI, Klasifikasi Teks.

## PENDAHULUAN

Arti perkembangan teknologi dan tingkat penggunaannya berdampak pada hampir semua aspek kehidupan manusia. Semua kegiatan yang dioperasionalkan membutuhkan teknologi itu sendiri. Istilah "digitalisasi" tampaknya menjadi populer dan dianggap sebagai alat untuk mempercepat pekerjaan manusia. Setiap orang merasakan kemudahan dan fleksibilitas yang diberikan oleh digitalisasi, yang mempermudah tugas dan aktivitas mereka. Hal ini juga berlaku untuk banyak organisasi, seperti bank (Mola et al., 2025).

Sebagaimana yang diketahui bank adalah suatu instansi yang berorientasi pada pengelolaan keuangan dari para nasabahnya. Nasabah adalah pelanggan aktif yang membutuhkan layanan perbankannya, dan setiap perbankan sangat memperhatikan kepuasan nasabah sebagai prioritas utama. Pasalnya hal tersebut juga menjadi daya tarik para nasabah untuk intens dan aktif menggunakan layanan perbankan yang menjadikan peningkatan produktivitas perbankan lebih optimal. Sehingga ketepatan dan efisiensi pelayanan bagi nasabah harus terus ditingkatkan untuk menjamin mutu pelayanan yang lebih baik (Sabrina Amanda Salsabila dkk, 2025).

Indonesia sendiri memiliki banyak sekali perusahaan perbankan yang besar dengan jumlah nasabah yang fantastis seperti (BRI), Bank Mandiri, (BCA), Bank Mega dan masih banyak lainnya. Dalam operasional sistemnya khususnya terkait perluasan jaringan tentunya harus menyesuaikan kapasitas dasar bank tersebut yaitu modal inti. Melalui cara tersebut maka bank tentunya memiliki kepercayaan penuh dalam sistem ketahanan yang baik sebab operasional kegiatannya cenderung pada kegiatan dan produk unggulan yang dimiliki (Lis Sintha Oppusunggu dan Yusuf Rombe M Allo, 2021).

Kajian ini memilih bank berdasarkan jumlah pelanggan terbanyak

dari masing-masing bank (Was'an, 2022). Menurut data terbaru, 2024, maka didapati data bahwa bank yang memiliki jumlah nasabah terbanyak salah satunya adalah Bank BNI dengan jumlah nasabah kurang lebih 64,03 nasabah (Muhajir dkk, n.d.). Sehingga dari data tersebut, peneliti memilih BNI sebagai subjek penelitian ini.

Sebagaimana yang dipahami bahwa variasi layanan dan bentuk layanan yang diberikan oleh setiap perbankan tentunya memunculkan ragam respon atau tanggapan dari para nasabahnya. Semisal ditemukan adanya kesalahan dari layanan yang diberikan maupun aktivitas transaksi di perbankan yang bersangkutan maka pihak nasabah akan melakukan kritik dan keluhan kepada bank yang dituju (Asmi Anugrah dan Rifqi Nur Wahyudi, 2025). Begitupun sebaliknya jika pelayanan yang didapat nasabah sangat memuaskan, efektif juga efisien tentu nasabah akan memberikan apresiasi positifnya. Sedangkan di masa sekarang ini perangkat digital memainkan peran besar dalam masyarakat. Perangkat digital menjadi alat untuk memberikan kritik dan ungkapan keluhannya (Muhammad Khoiril Insan dkk, 2023).

Melalui aplikasi Google Play Store, seorang nasabah dapat memberikan respon dan tanggapannya terkait layanan perbankan yang ia peroleh. Pendapat atau tanggapan nasabah yang disampaikan melalui platform Google Play Store menjadi objek kajian. Nasabah dapat mengungkapkan opini mereka secara langsung dan terbuka terhadap layanan perbankan dalam bentuk ulasan sebagai penilaian terhadap aplikasi mobile banking tersebut (Artanti Inez Tanggraeni dan Melkior NN Sitokdana, 2022). Seluruh tanggapan ini kemudian dianalisis berdasarkan sentimen positif maupun negatif, yang dapat menjadi masukan berharga bagi perusahaan perbankan sebagai bahan evaluasi dan dasar untuk meningkatkan mutu pelayanan kepada nasabah (Purnamasari, n.d.).

Analisis terhadap data ini dikenal

sebagai sentiment analysis atau opinion mining (Malebary, 2024). Sentiment analysis merupakan teknik dalam proses text mining, tujuan mengekstraksi, mengklasifikasikan, dan memahami perasaan atau pendapat yang terkandung dalam teks, seperti ulasan atau komentar pengguna. Salah satu metode yang paling umum digunakan dalam analisis perasaan adalah Support Vector Machine (SVM), yang dapat mengklasifikasikan data berdasarkan fitur tertentu. Metode ini mendukung keakuratan klasifikasi SVM dengan menggunakan solusi optimal global (Al, 2022). Pendekatan berbasis pembelajaran mendalam seperti *Bidirectional Encoder Representations from Transformers* (BERT) dapat memberikan hasil klasifikasi teks yang lebih akurat daripada SVM. BERT meningkatkan pemahaman pola sentimen yang lebih kompleks dengan menangkap konteks dua arah dalam teks (D. F. S. Dkk, 2024).

Dengan demikian, peneliti ingin menyelidiki analisis sentimen terhadap layanan perbankan, mengingat respon dan tanggapan nasabah terhadap layanan yang diberikan sangat berperan penting dalam mendorong kemajuan dan perkembangan bank yang bersangkutan. Penelitian ini menggunakan data dari ulasan Google Play Store yang menunjukkan perasaan positif dan negatif pengguna. Ada dua variabel dalam penelitian ini: variabel terikat adalah hasil klasifikasi sentimen, dan variabel bebas adalah teks ulasan dari kedua platform tersebut. Untuk melakukan klasifikasi sentimen, *Bidirectional Encoder Representations from Transformers* (BERT) dan *Support Vector Machine* (SVM) digunakan. Tujuan penggunaan kedua algoritma ini adalah untuk menentukan metode yang paling efektif untuk mengklasifikasikan sentimen pelanggan dengan membandingkan kinerja masing-masing.

## METODE

Fokus penelitian ini adalah untuk

mengevaluasi tingkat akurasi model klasifikasi sentimen dan menentukan kecenderungan sentimen positif, negatif, atau netral yang muncul. Fokus penelitian ini adalah layanan perbankan BNI, dan sumber datanya adalah ulasan pengguna aplikasi yang dikumpulkan dari Google Play Store melalui teknik web scraping. Untuk membuat laporan penelitian, peneliti menggunakan Microsoft Word 2021, sedangkan Mendeley Desktop digunakan untuk mengelola dan mengumpulkan referensi.

Kajian ini melihat komentar pemakai tentang aplikasi Wondr milik BNI, tersedia di Google Play Store, dan menggambarkan pengalaman dan tanggapan pengguna tentang layanan perbankan BNI. Ulasan tersebut mencakup pendapat positif, negatif, dan netral. Sampel penelitian diambil dari sebagian ulasan tersebut untuk merepresentasikan keseluruhan populasi, dengan Pengumpulan data dilakukan dengan teknik web scraping berbasis Python.

Untuk memulai proses pelatihan model SVM, teks ulasan diubah menjadi representasi numerik TF-IDF. Proses pelatihan menggunakan representasi ini untuk membagi sentimen ulasan ke dalam kategori positif, netral, dan negatif. Sedangkan untuk memproses teks berbahasa Indonesia, Pelatihan Model Indobert menggunakan tokenizer indobert-base-p1. Oleh karena itu, hasilnya adalah label emosi positif, netral, dan negatif.

## HASIL DAN PEMBAHASAN

### Implementasi IndoBERT

Dalam penelitian ini, model IndoBERT-base-p1 digunakan untuk menganalisis aplikasi wondr BNI. Model ini, yang telah dilatih sebelumnya dan berbasis arsitektur Transformer, dapat memproses Bahasa Indonesia dan menangkap konteks dua arah (bidirectional) dari kalimat. Berbeda dengan metode tradisional seperti SVM

(Support Vector Machine) yang mengandalkan statistik kaku, IndoBERT menggunakan mekanisme *self-attention* untuk memahami nuansa bahasa yang lebih kompleks, termasuk penggunaan singkatan, bahasa gaul, dan struktur kalimat tidak resmi yang sering muncul dalam feedback nasabah.

```
import os
os.environ["WANDB_DISABLED"] = "true"

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, accuracy_score, confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt
import torch
from transformers import BertTokenizer, BertForSequenceClassification, Trainer, TrainingArguments
from transformers import DataCollatorWithPadding
from datasets import Dataset

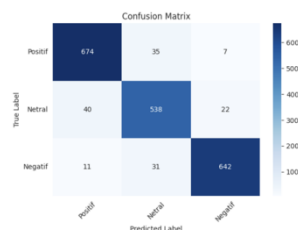
# Load data
df = pd.read_csv("Hasil_labelling_data.csv")
df.info()
df.head(5)
```

**Gambar 1 Code Install IndoBERT**

Proses pengerjaan teknis dilakukan secara sistematis dengan mengacu pada alur kode dibawah ini (Sneha Leela Jacob and Parveen Sultana Habibullah, 2025):

1. Instalasi Perangkat Lunak: Tahap awal dilakukan dengan memasang pustaka yang diperlukan melalui perintah `!pip install transformers datasets scikit-learn seaborn matplotlib` agar seluruh fungsi analisis tersedia dalam lingkungan kerja.
2. Tokenisasi Teks: Untuk memproses ulasan, digunakan BertTokenizer dari varian IndoBERT yang menerapkan metode *subword tokenization* (WordPiece). Teknik ini sangat efektif untuk menangani kata-kata tak ada pada kamus (*Out-of-Vocabulary*), sehingga tidak ada informasi penting yang terbuang.
3. Pengaturan Panjang Kalimat: Batas maksimal urutan teks (*maximum sequence length*) ditetapkan pada angka 128 token. Penentuan ini didasarkan pada karakteristik ulasan pengguna wondr by BNI yang umumnya cukup singkat (di bawah 100 kata), sehingga proses komputasi menjadi lebih efisien tanpa memotong isi ulasan yang ada.

### Perhitungan Confusion Matrix



**Gambar 2 Confusion Matrix IndoBERT**

Berdasarkan pemetaan *confusion matrix* 3×3, model IndoBERT terbukti memiliki reliabilitas tinggi dengan akurasi total mencapai 93%. Keunggulan utama terlihat pada klasifikasi sentimen negatif yang mencatatkan presisi 0,96, di mana 642 data berhasil diprediksi secara akurat dengan risiko kesalahan sangat minim. Pada kategori positif, model menunjukkan sensitivitas tajam dengan skor *recall* 0,94, yang berarti hanya sekitar 6% ulasan apresiatif yang gagal terdeteksi.

Meskipun kelas netral menghadapi tantangan ambiguitas kata kunci yang menyebabkan 57 ulasan salah terklasifikasi, model tetap stabil dengan *f1-score* 0,89. Sinkronisasi antara perhitungan manual dan otomatis ini mengonfirmasi bahwa IndoBERT mampu menggeneralisasi pola bahasa ulasan wondr by BNI secara adil dan konsisten di seluruh kategori sentimen.

Sebagai bentuk validasi mendalam, peneliti melakukan penghitungan manual melalui *Confusion Matrix* pada 2.000 data uji untuk menjamin transparansi performa model di setiap kategori sentimen aplikasi wondr by BNI. Hasilnya menunjukkan akurasi yang sangat solid: pada kategori positif, model berhasil mengidentifikasi 674 dari 716 data aktual, sementara pada kategori negatif, ketajaman model terbukti dengan keberhasilan mengklasifikasikan 642 dari 684 ulasan secara tepat. Meski kelas netral menghadapi tantangan ambiguitas dengan 538 deteksi akurat dari 600 data, margin kesalahan yang rendah di seluruh kelas menegaskan bahwa IndoBERT memiliki konsistensi tinggi dalam membedakan nuansa emosi pengguna secara objektif dan reliabel.

**Tabel 1 Hasil Classification Report dari Confusion Matrix Model IndoBERT**

Kelas	Perfor ma		
	<i>Precision</i>	<i>Recall</i>	<i>F1- Score</i>
<b>Negatif</b>	0.96	0.94	0.95
<b>Positif</b>	0.93	0.94	0.94
<b>Netral</b>	0.89	0.90	0.89
<b>Akurasi</b>	0.93		

Tabel 4.7.6 Hasil Laporan Klasifikasi Model IndoBERT Confusion Matrix menunjukkan bagaimana evaluasi model berhasil mengkategorikan sentimen menjadi negatif, positif, dan netral. Secara keseluruhan, model IndoBERT mencapai nilai akurasi sebesar 0,93 (93%), menunjukkan bahwa sebagian besar data uji diklasifikasikan dengan tepat. Pada kategori sentimen negatif, model menunjukkan performa paling unggul dengan nilai Precision 0,96, Recall 0,94, dan F1-Score 0,95. Hasil ini menunjukkan bahwa IndoBERT sangat andal dalam mengidentifikasi ulasan bernada negatif, dengan tingkat kesalahan prediksi yang relatif rendah. Nilai presisi yang tinggi menunjukkan bahwa sebagian besar prediksi negatif yang dihasilkan model memang benar, sementara nilai recall yang tinggi menunjukkan kemampuan model untuk mengumpulkan hampir seluruh data negatif yang ada.

Selain itu, kategori sentimen positif memiliki kinerja yang sangat baik, dengan F1-Score sebesar 0,88, Precision sebesar 0,89, dan Recall sebesar 0,94. Ini menunjukkan bahwa model tidak hanya mampu mengenali ulasan positif secara akurat, tetapi juga memiliki konsistensi yang baik antara cakupan data positif yang berhasil ditemukan dan ketepatan prediksi. Hasil menunjukkan bahwa IndoBERT mampu menangkap karakteristik sentimen netral dengan baik, meskipun nilai performanya sedikit lebih rendah dibandingkan dua kategori lainnya. Adanya tumpang tindih makna antara sentimen netral dengan sentimen positif dan negatif dapat menyebabkan

perbedaan dalam kinerja ini, seperti yang ditunjukkan oleh matriks confusion..

Secara keseluruhan, evaluasi ini menunjukkan bahwa model IndoBERT memiliki kestabilan dan keandalan yang tinggi, dengan semua metrik evaluasi berada di atas 0,89. Pemetaan confusion matrix 3×3 memperlihatkan bahwa model ini mampu membedakan antara sentimen positif, negatif, dan netral dengan tingkat kesalahan klasifikasi yang sangat minim. BNI menganggap IndoBERT sebagai sarana yang sangat tepat untuk menganalisis sentimen ulasan pengguna aplikasi Wondr, dengan tingkat akurasi mencapai 93%. Penggunaan model ini memberikan keuntungan besar dibandingkan model konvensional seperti SVM dalam hal pemahaman makna. IndoBERT mampu menghasilkan representasi *embedding* yang jauh lebih mendalam dan kaya akan konteks dibandingkan metode *bag-of-words*. Hal ini secara langsung berdampak pada tingkat akurasi yang lebih tinggi saat memisahkan komentar Positif, Negatif, maupun Netral. Melalui integrasi fungsi Trainer dan TrainingArguments seperti yang terlihat pada kode, model dapat dilatih untuk mengenali pola kepuasan atau keluhan nasabah dengan sangat presisi.

### Implementasi Model SVM

Untuk memisahkan data berdasarkan ruang fitur ke dalam kelas sentimen yang berbeda, algoritma klasifikasi supervised learning Model Support Vector Machine (SVM) mencari hyperplane optimal. SVM dipilih karena kemampuan untuk menangani data yang sangat besar dan kompleks, yang merupakan karakteristik umum data teks ulasan (Zain, 2025).

Implementasi model dimulai dengan proses impor library penelitian utama, seperti pandas dan numpy untuk pengolahan data, scikit-learn untuk pemodelan pembelajaran mesin, dan matplotlib dan seaborn untuk tampilan hasil evaluasi. Hasil pelabelan sentimen sebelumnya kemudian dimasukkan ke

dalam dataset. Metode Term Frequency–Inverse Document Frequency (TF-IDF) digunakan dalam penelitian ini, yang menggunakan *TfidfVectorizer* pada tahap ekstraksi fitur (H. B. Dkk, 2025). Dengan memprioritaskan kata-kata diskriminatif dan mengurangi pengaruh kata-kata umum, metode ini berhasil mengubah data teks menjadi representasi numerik. Karena kemampuan untuk menunjukkan kepentingan relatif suatu kata dalam dokumen, TF-IDF dinilai efektif dalam analisis sentimen teks pendek, seperti ulasan pengguna aplikasi.

Setelah itu, klasifikasi dilakukan dengan menggunakan Support Vector Classifier (SVC). Model SVM ini dipilih karena tidak membutuhkan sumber daya komputasi yang kuat seperti GPU, tetapi masih mampu menghasilkan kinerja yang stabil dan akurat pada dataset berukuran menengah. Untuk melihat distribusi kesalahan klasifikasi, Akurasi, ketepatan, recall, dan skor f1 adalah metrik yang digunakan untuk mengevaluasi model. Selain itu, model divisualisasikan melalui matriks confusion.

**Tabel 2 Hasil Classification Report dari Confusion Matrix Model SVM**

Kelas	Performa		
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Negatif	0.91	0.94	0.93
Positif	0.92	0.95	0.94
Netral	0.90	0.83	0.86
Akurasi	0.91		

Dengan tingkat performa yang relatif tinggi, Dalam Tabel 4.8.6, hasil laporannya menunjukkan bahwa model Support Vector Machine (SVM) mampu mengkategorikan sentimen ke dalam tiga kategori: negatif, positif, dan netral. Secara keseluruhan, model SVM memperoleh nilai akurasi sebesar 0,91, atau 91%, yang menandakan bahwa mayoritas data uji diprediksi dengan tepat. Nilai akurasi 0,91 untuk kategori sentimen negatif menunjukkan bahwa

model mampu mengidentifikasi ulasan bernada negatif dengan baik dan dapat mengelola sebagian besar data negatif.

Kinerja terbaik ditunjukkan oleh kelas sentimen positif; mereka memiliki nilai precision 0,92, recall 0,95, dan F1-score 0,94. Nilai precision yang tinggi menunjukkan kemampuan model SVM untuk mengidentifikasi ulasan positif, sementara nilai precision yang tinggi menunjukkan ketepatan prediksi yang konsisten..

Sementara itu, model menemukan nilai precision sebesar 0,90, recall sebesar 0,83, dan F1-score sebesar 0,86. Nilai recall yang lebih rendah dibandingkan dengan kelas lainnya menunjukkan bahwa sebagian besar data sentimen netral masih cenderung termasuk dalam kelas positif atau negatif. Ini mungkin karena karakteristik sentimen netral sering memiliki kemiripan konteks dengan dua kelas lainnya.

Secara keseluruhan, hasil evaluasi melalui confusion matrix menunjukkan bahwa model SVM memiliki kinerja yang stabil dan seimbang di seluruh kelas sentimen. Dengan nilai precision, recall, dan F1-score yang sebagian besar di atas 0,85, model SVM dinilai mampu melakukan klasifikasi sentimen dengan efektif dan layak digunakan dalam analisis ulasan pengguna aplikasi.

### Perbandingan Kinerja Model IndoBERT dan SVM

Setelah seluruh tahapan pengujian terhadap model IndoBERT dan *Support Vector Machine* (SVM) diselesaikan, tahap selanjutnya adalah menganalisis dan membandingkan kinerja kedua model dalam mengkategorikan sentimen ulasan pengguna dari aplikasi Wondr milik BNI. Perbandingan ini dilakukan untuk menentukan model untuk membagi sentimen ulasan pengguna dari aplikasi Wondr milik BNI. Perbandingan ini dilakukan untuk menentukan model mana yang memberikan hasil terbaik dalam memahami dan mengkategorikan sentimen ulasan ke dalam kategori negatif, netral, dan positif.

**Tabel 3**Perbandingan Hasil Akurasi Model IndoBERT dan SVM

Metrik	SV M	BER T	Selisih (%)	Mo del Terb aik
Accura cy	0.83 287 3	0.92 7000	9.41	IndoB ERT
Precisi on	0.83 227 4	0.92 7254	9.50	IndoB ERT
Recall	0.83 287 3	0.92 7000	9.41	IndoB ERT
F1- Score	0.83 096 0	0.92 7081	9.61	IndoB ERT

Seperti yang ditunjukkan dalam Tabel 4.9, kinerja model SVM dan IndoBERT sangat berbeda pada semua metrik evaluasi yang digunakan. IndoBERT memiliki nilai akurasi, ketepatan, recall, dan nilai F1 yang lebih tinggi dibandingkan dengan SVM. Selisih performa kedua model adalah 9% hingga 10%, menunjukkan bahwa IndoBERT memiliki kemampuan klasifikasi yang lebih baik; SVM hanya memiliki nilai 0,832873. Situasi serupa juga terlihat dengan nilai precision dan recall. IndoBERT dapat mengklasifikasikan sentimen dengan lebih tepat dan lengkap daripada SVM, dan nilai F1-nya yang lebih tinggi menunjukkan bahwa model ini memiliki keseimbangan yang lebih baik antara precision dan recall.

Secara keseluruhan, hasil perbandingan menunjukkan bahwa IndoBERT adalah model yang paling cocok untuk penelitian ini. Ini mungkin karena keunggulannya dalam memahami konteks dan makna kata secara lebih mendalam, yang membuatnya lebih baik daripada metode SVM konvensional dalam menganalisis sentimen.

## SIMPULAN

Hasil penelitian menunjukkan

bahwa (1) model Support Vector Machine (SVM) masih gagal mengklasifikasikan sentimen ulasan pengguna aplikasi Wondr oleh BNI, dengan akurasi hanya 0,832873. (2) Representations Bidirectional Encoder dari Transformers menunjukkan kinerja yang lebih baik untuk semua metrik evaluasi yang diuji, yaitu akurasi. (3) BERT lebih baik dalam mengklasifikasikan data secara proporsional dan akurat. BERT memiliki keunggulan dalam hal ketepatan prediksi (accuracy) dan kemampuan untuk mengurangi kesalahan klasifikasi, yang ditunjukkan oleh nilai precision dan recall yang lebih tinggi. Selain itu, nilai F1-score yang lebih tinggi BERT menunjukkan bahwa model ini lebih baik dalam mengimbangi precision dan recall daripada SVM. Kesimpulan harus mengindikasikan secara jelas hasil-hasil yang diperoleh, kelebihan dan kekurangannya, serta kemungkinan pengembangan selanjutnya.

## DAFTAR PUSTAKA

- Al, X. L. et. (2022). Social network sentiment classification method combined Chinese text syntax with graph convolutional neural network. *Egyptian Informatics Journals*, 23(1).
- Artanti Inez Tanggraeni dan Melkior NN Sitokdana. (2022). *Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naive Bayes*. 9(2), 785–795.
- Asmi Anugrah dan Rifqi Nur Wahyudi. (2025). Pengaruh Digital Banking Dan Efisiensi Operasional Terhadap Kinerja Pelayanan Bank Sulselbar Cabang Antang. *Jurnal Axegnal*, 2(1).
- Dkk, D. F. S. (2024). Analisis Sentimen Film Dirty Vote Menggunakan BERT (Bidirectional Encoder Representations from Transformers). *Jurnal JTJK*, 8(2).
- Dkk, H. B. (2025). Sentiment Classification Analysis of Tokopedia

- Reviews Using TF-IDF, SMOTE, and Traditional Machine Learning Models. *Journal of Applied Informatics and Computing (JAIC)*, 9(5).
- Lis Sintha Oppusunggu dan Yusuf Rombe M Allo. (2021). *Kecukupan Modal Inti Bank*. Widina Bhakti Persada.
- Malebary, S. J. and A. W. A. (2024). A Stacking Ensemble Based on Lexicon and Machine Learning Methods for the Sentiment Analysis of Tweets. *Mathematics*, 12(21). <https://doi.org/https://doi.org/10.339>
- Mola, S. A. S., Putri, D., Polly, N., & Rumlaklak, N. D. (2025). *Sentiment Analysis on User Reviews of the Edlink Application Using the Random Forest Classifier Method*. 15(1), 20–27.
- Muhajir dkk. (n.d.). *Analisis Laporan Keuangan Bank Negara Indonesia dan Bank Negara Indonesia Syariah*.
- Muhammad Khoirul Insan dkk. (2023). ANALISIS SENTIMEN APLIKASI BRIMO PADA ULASAN PENGGUNA DI GOOGLE PLAY MENGGUNAKAN ALGORITMA NAIVE BAYES. *JATI*, 7(1).
- Purnamasari, D. (n.d.). *Pengantar Metode Analisis Sentimen*.
- Sabrina Amanda Salsabila dkk. (2025). Komparasi Kinerja Model Naive Bayes, SVM, Dan Bert Dalam Klasifikasi Sentimen Ulasan Pada Aplikasi YUMMY. *Storage*, 4(2).
- Sneha Leela Jacob and Parveen Sultana Habibullah. (2025). A Systematic Analysis and Review on Intrusion Detection Systems Using Machine Learning and Deep Learning Algorithms. *Journal of Computational Cognitive Engineering*, 4(2).
- Was'an, G. H. (2022). Pendekatan Analisis Vector Error Correction Model (VECM) Dalam Hubungan Kondisi Makro Ekonomi Dengan Non Performing Financing Berdasarkan Pengelompokan Modal Inti Bank Umum Syariah di Indonesia. *Jurnal Neraca Peradaban*, 2(2).
- Zain, A. F. dkk. (2025). Analisis Sentimen Ulasan Pengguna iPhone dengan Pendekatan Hibrida RoBERTa dan XGBoost. *Jurnal Algoritma*, 22(1).