

Penerapan Pendekatan *Adjustment Factor* dalam Analisis Sentimen Kepuasan Konsumen Menggunakan Indo-BERT

Application of the Adjustment Factor Approach in Consumer Satisfaction Sentiment Analysis Using Indo-BERT

Astrid Noviana Paradhita^{a,1,*}, Myrtana Pusparisti^{a,2}, Putra Pamungkas^{b,3}

^aManajemen Bisnis, Universitas Sebelas Maret, Surakarta, Indonesia

^bBisnis Digital, Universitas Sebelas Maret, Surakarta, Indonesia

¹astrid.noviana@staff.uns.ac.id; ²myrta@staff.uns.ac.id; ³putra.pamungkas@staff.uns.ac.id

*corresponding author

Informasi Artikel	ABSTRAK
Diserahkan : 6 Maret 2026 Diterima : 18 Mei 2026 Direvisi : 4 Juni 2026 Diterbitkan : 24 Agustus 2026 Kata Kunci: Analisis Sentimen TF-IDF SVM Indo-BERT <i>Adjustment Factor</i>	Pengukuran tingkat kepuasan konsumen menjadi komponen penting dalam meningkatkan kualitas layanan dan pengambilan keputusan dalam suatu jasa manufaktur yang memiliki kewajiban untuk terus meningkatkan kualitas layanan dan fungsi terhadap kualitas jasa yang diberikan. Data umpan balik konsumen yang umumnya berbentuk teks dan tidak terstruktur membuat analisis sulit dilakukan. Oleh karena itu, diperlukan pendekatan analisis khusus yang mampu memahami konteks bahasa Indonesia secara komputasional. Penelitian ini bertujuan untuk menganalisis sentimen kepuasan konsumen terhadap kualitas layanan jasa yang diberikan. Analisis sentimen juga dilakukan untuk menangkap suasana teks yang disesuaikan dengan frasa tertentu dalam kualitas jasa yang memiliki tingkat sensitivitas lebih kuat. Dengan mengintegrasikan TF-IDF, SVM, Indo-BERT dan <i>rule-based adjustment factor</i>, penelitian ini berhasil melakukan analisis sentimen terhadap 10.421 data umpan balik konsumen. Hasil penelitian menunjukkan bahwa integrasi metode dapat meningkatkan konsistensi dalam mendeteksi sentimen konsumen dan menghasilkan klasifikasi yang lebih representatif terhadap konteks kualitas layanan jasa. Analisis sentimen menghasilkan performa yang baik dengan tingkat akurasi mencapai 90% dan tingkat presisi mencapai 85.7%. Selain itu, juga diperoleh nilai <i>recall</i> sebesar 100% dan F1-score 92.3%.
Keywords: Sentiment Analysis TF-IDF SVM Indo-BERT <i>Adjustment Factor</i>	ABSTRACT <i>Measuring customer satisfaction is crucial for enhancing service quality and informing decision-making within service-oriented manufacturing firms, which are obligated to improve and oversee service quality continuously. Customer feedback data—typically unstructured and text-based—poses challenges for analysis; therefore, a specialized analytical approach capable of computationally understanding the context of the Indonesian language is required. This study aims to analyze customer satisfaction sentiment regarding the services provided. Sentiment analysis was also conducted to capture the textual tone, specifically accounting for certain phrases that carry heightened sensitivity. By integrating TF-IDF, SVM, Indo-BERT, and a rule-based adjustment factor, the study successfully analyzed sentiment in 10,421 customer feedback entries. The results demonstrate that this integrated approach enhances consistency in detecting customer sentiment and yields classifications that better reflect the context of services quality. The sentiment analysis achieved strong performance, with an accuracy of 90%, precision of 85.7%, recall of 100%, and an F1-score of 92.3%.</i>
This is an open access article under the CC-BY-SA license. 	

I. Pendahuluan

Mengukur tingkat kepuasan konsumen sangat diperlukan untuk memperoleh umpan balik dan input sebagai bagian dari upaya dalam meningkatkan strategi pelayanan [1]. [2] menyatakan metode khusus diperlukan untuk meneliti komentar, umpan balik, jawaban, dan rekomendasi dari pengguna layanan. Hal ini dikarenakan data dan informasi yang tersedia umumnya berbentuk tidak terstruktur berupa kalimat maupun opini. Analisis sentimen merupakan metode yang dapat digunakan untuk mengetahui tingkat kepuasan konsumen tidak terkecuali dalam kualitas pelayanan jasa di perusahaan X.

Dalam perkembangan ilmu pengetahuan dan teknologi, transformasi pemanfaatan teknologi telah digunakan dalam skala besar (*big data*). Di berbagai sektor organisasi baik pemerintahan, bisnis, maupun pendidikan, konsumen memiliki akses yang luas sehingga mampu memfasilitasi interaksi, opini, dan umpan balik secara *realtime*. Fenomena ini menghasilkan volume data teks yang besar dari ulasan pengguna layanan akhir terhadap suatu kebijakan maupun pelayanan. Analisis dari data teks tersebut menjadi komponen penting dalam memahami sentimen dan preferensi konsumen [3][4].

Dalam beberapa tahun terakhir, metode analisis sentimen dengan *Natural Language Processing* (NLP) berbasis *transformer* telah menjadi tonggak penting dalam pengolahan bahasa alami. Model *pre-trained* seperti BERT (*Bidirectional Encoder Representations from Transformers*) mampu memahami konteks bahasa secara mendalam. BERT dapat mengkategorisasikan teks dalam konsep *bag of words* (BoW) [5]. Namun demikian, BERT hanya memiliki performa yang optimal dalam menelaah teks berbahasa Inggris [6]. Oleh karena itu, penerapan Indo-BERT dalam analisis sentimen diperlukan untuk mengatasi keterbatasan sumber daya yang kurang terlayani dari model bahasa Inggris [7]. Indo-BERT menjadi pilihan model yang tepat karena terbukti dapat meningkatkan akurasi klasifikasi dan interpretabilitas data teks bahasa Indonesia dibandingkan metode BERT [8]. [9] juga membuktikan bahwa Indo-BERT menunjukkan performa yang baik dalam mengklasifikasi sentimen bahasa Indonesia ke dalam kategori positif, negatif, dan netral, serta mampu mengidentifikasi area sentimen.

Meski demikian, tantangan signifikan dalam analisis sentimen masih muncul ketika model analisis sentimen harus menangkap nuansa bahasa yang kompleks seperti sarkasme, ironi, dan nilai sentimen yang sangat kontekstual dari konsumen. Selain itu, kesulitan juga muncul dari suatu opini yang tidak mengandung emosi maupun preferensi yang diperlukan. Sebagai respon atas keterbatasan ini, penelitian ini akan mengeksplorasi mekanisme penyesuaian (*adjustment factor*). [10] menyatakan bahwa pendekatan *judgemental adjustment* diperlukan untuk meningkatkan akurasi sentimen dan harus dilakukan secara objektif berdasarkan kondisi khusus yang dapat diberlakukan secara general untuk seluruh data.

Pendekatan ini dapat dilakukan dengan menerapkan beberapa aturan dalam analisis sentimen dalam bentuk *rule-based adjustment factor*. Pembuatan *rule-based* yang sesuai dengan pola dalam BoW seringkali terbatas. Jika, *rule-based* berlebihan dapat menyebabkan kesulitan dalam penyesuaian parameter dan meningkatkan kompleksitas aturan [11]. Oleh karena itu, *rule-based* yang dibuat juga harus disesuaikan dengan kondisi nyata dalam konteks kalimat yang sudah ada dalam kamus *pre-training* BoW [12].

Merujuk pada penelitian sebelumnya [13] melakukan analisis sentimen terhadap serangan siber di BSI yang menyebabkan hilangnya data dari 15 juta nasabah dan 1.5 terabyte data internal menggunakan Indo-BERT. Hasil penelitian menunjukkan adanya permasalahan pada proses transaksi di mobile banking BSI yang lama sehingga perlu dilakukan *upgrade* teknologi. Sedangkan, pada penelitian lain, [14] menyampaikan bahwa penggunaan Indo-BERT masih perlu ditingkatkan untuk mendapatkan hasil skoring analisis sentimen berbasis aspek. Penelitian ini menunjukkan bahwa proses deteksi sentimen dan preferensi terhadap suatu kalimat maupun opini memerlukan proses identifikasi yang lebih dalam untuk mendapatkan kunci skoring berdasarkan konteks spesifik yang dimiliki pada suatu domain sentimen tertentu.

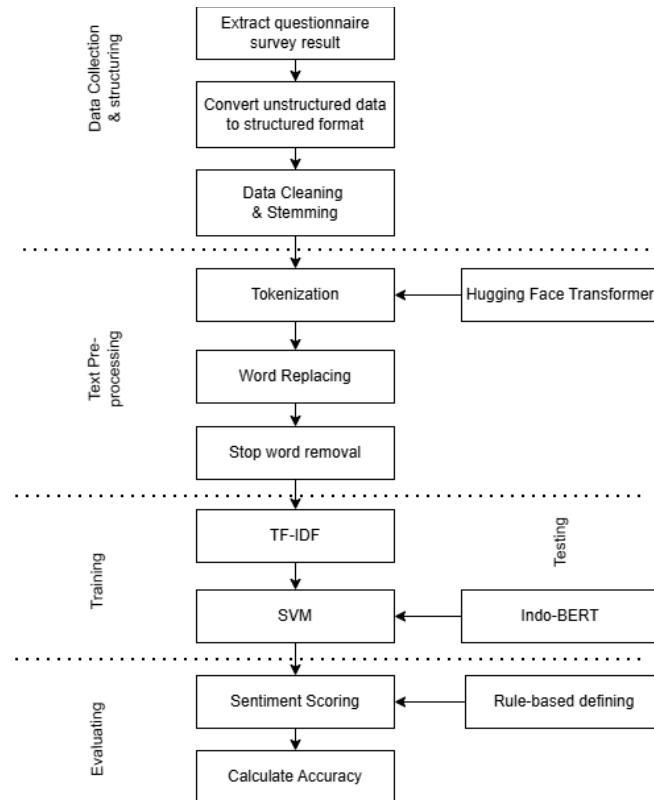
Penelitian ini akan menerapkan dan mengevaluasi penggunaan *rule-based adjustment factor* untuk memperkaya fitur representasi dan kalibrasi skor sentimen. Pendekatan *adjustment factor* merupakan bentuk penyesuaian berbasis domain terhadap bobot kata dengan mempertimbangkan distribusi opini konsumen secara holistik. Dengan demikian, analisis sentimen bukan hanya menghasilkan label polaritas sentimen, namun juga refleksi kualitas umpan balik yang lebih akurat dan bermakna dalam konteks pengukuran kepuasan konsumen.

II. Metode

Penelitian ini terlaksana dengan mengadopsi pendekatan kuantitatif berbasis konteks sentimen konsumen dalam mengkategorisasikan kalimat atau opini. Penelitian dilakukan dengan mengintegrasikan teknik NLP berbasis arsitektur *Transformer* untuk menganalisis sentimen kepuasan konsumen secara komprehensif. Penelitian juga memasukkan teknik kontekstualisasi dengan memberlakukan *rule-based adjustment factor* untuk menangkap nuansa kontekstual bahasa Indonesia. Seluruh penelitian dilakukan menggunakan Python pada platform Google Collaboratory melalui alur penelitian yang sistematis.

A. Alur Penelitian

Penelitian ini dilakukan berdasarkan alur kerja yang sudah ditetapkan. Gambar 1 menunjukkan struktur alur penelitian analisis sentimen konsumen.



Gambar 1. Alur Penelitian

B. Data Collection & Structuring

Tahap awal penelitian merupakan fondasi penelitian yang dimulai dengan mengekstraksi hasil survei kuesioner. Data yang diperoleh merupakan opini tekstual dari konsumen dalam bentuk kalimat bebas berupa kritik dan saran atas layanan maupun kebijakan yang berlaku [15]. Tahapan ini mencerminkan praktik *data driven research*, di mana kualitas struktur data menjadi komponen penting yang menentukan performa model analitik sentimen [16].

C. Text Preprocessing

Tahap kedua adalah *text preprocessing* yang dirancang untuk menyiapkan data teks agar sesuai dengan kebutuhan paradigma analisis [17]. Pada tahap ini peneliti menggunakan dua jenis pendekatan analisis melalui pendekatan statistik klasik dan pendekatan representasi kontekstual. Kebaharuan terletak pada integrasi Hugging face Transformer yang secara langsung digunakan pada saat tokenisasi. Tokenisasi menggunakan Hugging face Transformer memungkinkan pemecahan teks berbasis *sub-word unit* yang sangat efektif untuk menangani kosakata luar (*out of vocabulary*) [14]. Pada tahap ini juga dilakukan proses *word replacing* untuk menggantikan dan menyesuaikan kata yang tidak baku seperti penggunaan singkatan guna meningkatkan konsistensi semantik data [18]. Selain itu, ada juga proses *stop word removal* untuk menghilangkan kata-kata yang tidak berkontribusi signifikan terhadap pembentukan sentimen. Proses ini juga dapat mengurangi dimensi data dan noise yang dapat menurunkan tingkat akurasi analisis sentimen [19].

D. Training & Testing

Tahap ketiga dalam penelitian adalah *training & testing* model dengan menerapkan dua pendekatan representasi perkembangan teknologi analisis teks. Pendekatan pertama menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* sebagai teknik ekstraksi fitur berbasis statistik. Untuk mendapatkan nilai TF-IDF perlu dilakukan tiga jenis perhitungan yaitu [20]:

Perhitungan nilai TF dengan persamaan 1.

$$TF(t, d) = \frac{\text{Total number of term in the document } d}{\text{Number of occurrence of term } t \text{ in the document } d} \quad (1)$$

Melakukan perhitungan nilai IDF dapat menggunakan persamaan 2.

$$IDF(t, D) = \left(\frac{\text{Number of document containing the term } t+1}{\text{Total number of documents in the corpus } |D|} \right) + 1 \quad (2)$$

Sehingga untuk mendapatkan nilai TF-IDF digunakanlah persamaan 3.

$$TF - IDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (3)$$

Pendekatan selanjutnya yaitu penerapan Support Vector Machine (SVM) untuk membangun model klasifikasi sentimen. SVM digunakan karena memiliki performa yang stabil pada data teks berdimensi tinggi [2]. Sedangkan, Indo-BERT secara khusus dilatih pada korpus bahasa Indonesia untuk menangkap konteks semantik dan sintaksis opini secara lebih mendalam [21]. Dalam analisis sentimen berbasis NLP, penelitian ini mengimplementasikan model Indo-BERT menggunakan *pretrained weight* dari Hugging Face dengan varian Indo-BERT Base-P1. Proses fine-tuning selanjutnya dilakukan pada data sentimen konsumen menggunakan *framework Transformer* dari Hugging Face pada *platform* Google Collaboratory. Parameter pelatihan yang digunakan meliputi *learning rate* sebesar 2×10^{-5} *batch size* 16, dan jumlah *epoch* sebanyak lima iterasi pelatihan. Pada tahap ini, data dibagi menjadi data pelatihan dan data pengujian untuk mengukur kemampuan generalisasi model dalam mengenali polaritas sentimen konsumen. Perpaduan SVM dan Indo-BERT menjadi krusial untuk menilai kehandalan model analisis sentimen dalam mengukur kepuasan konsumen yang dinamis dan kontekstual.

E. Evaluasi

Tahap terakhir yaitu evaluasi model. Berbeda dari penelitian analisis konvensional yang berhenti pada klasifikasi polaritas sentimen, penelitian ini melangkah lebih lanjut dengan diberlakukannya *rule-based adjustment factor* pada saat proses skoring dilakukan. Setiap hasil sentimen yang sudah dikonversi menjadi skor numerik untuk menunjukkan tingkat intensitas kepuasan konsumen. Hasil pendalaman inilah yang berikutnya menjelma menjadi langkah baru untuk mempertimbangkan konteks domain, bobot kata kunci strategis, dan karakteristik umpan balik yang didapatkan. Integrasi TF-IDF, SVM, Indo-BERT dengan *rule-based adjustment factor* menghasilkan kalibrasi model berbasis AI yang lebih selaras dengan realitas pengambilan keputusan organisasi, sehingga dapat menjembatani kesenjangan antara kecerdasan mesin dan interpretasi manusia [11].

Pada tahap akhir, peneliti juga melakukan evaluasi performa model analisis sentimen. Hasil evaluasi digunakan untuk mengetahui efektifitas metode TF-IDF dan SVM dengan Indo-BERT dan *rule-based adjustment factor* untuk menegaskan peningkatan kualitas analisis sentimen kepuasan konsumen. Secara algoritmik, pemberian nilai *adjustment factor* ditentukan melalui *expert judgement* yang melibatkan pihak dari perusahaan X dan Tim Ahli dari penelitian. [22] menyatakan bahwa perhitungan nilai performa bisa dilakukan dengan menghitung nilai akurasi dengan persamaan (4), nilai presisi dengan persamaan (5), nilai recall dengan persamaan (6), dan F1-score dengan persamaan (7).

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

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (6)$$

$$F1 - \text{score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (7)$$

Keterangan:

TP (True Positive): Jumlah prediksi positif dan *ground truth* positif

TN (True Negative): Jumlah prediksi negatif dan *ground truth* juga negatif

FP (False Positive): Jumlah prediksi positif namun *ground truth* negatif.

FN (False Negative): Jumlah prediksi negatif, namun *ground truth* positif.

Dalam implementasinya, *ground truth* didasarkan pada skala likert survei yang juga dilakukan untuk mendapatkan nilai kepuasan konsumen secara kuantitatif. Survei dilakukan terhadap konsumen dengan skala likert 1 – 6. Tabel 1 menunjukkan skala likert yang digunakan sebagai basis *ground truth* dalam penelitian.

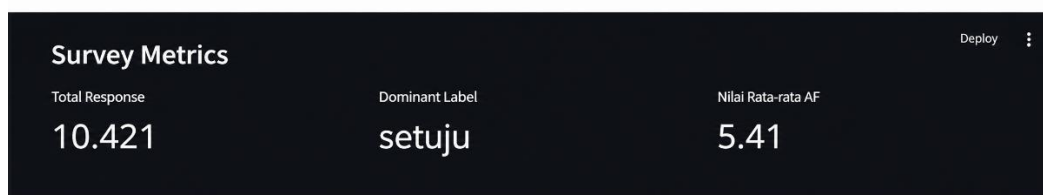
Tabel 1. Skala likert survei

Skala	Keterangan
1	Sangat Tidak Setuju
2	Tidak Setuju
3	Kurang Setuju
4	Cukup Setuju
5	Setuju
6	Sangat Setuju

III. Hasil dan Pembahasan

A. Data Collection & Structuring

Tahap awal penelitian dilakukan dengan mengumpulkan data berupa hasil kuesioner survei kepuasan konsumen. Survei dilakukan dari tanggal 8 November 2025 – 31 Desember 2025. Hasil survei berisi komentar atau umpan balik dalam bentuk teks bebas. Total data umpan balik yang digunakan dalam analisis sentimen di penelitian ini adalah 10.421 data responden. Gambar 2 menunjukkan hasil pengisian umpan balik oleh konsumen pada sistem survei yang telah dikembangkan.



Gambar 2. Data Survei Kepuasan *Konsumen*

Dari Gambar 2 peneliti mendapatkan informasi terkait jumlah konsumen yang telah mengisi survei. Seluruh data yang diperoleh pada tahap ini belum terstruktur sehingga perlu dikonversi menjadi format yang lebih mudah untuk dianalisis secara komputasional. Setelah proses ekstraksi data dan penyusunan format data tabular, didapatkan informasi seperti id responden, waktu pengisian respon, dan teks umpan balik yang diberikan oleh responden. Data lain yang tidak berpengaruh terhadap hasil analisis sentimen tidak dimasukkan dalam tahap *teks preprocessing* dan hanya digunakan untuk melakukan *filtering data* sesuai kebutuhan. Tabel 2 menunjukkan contoh hasil ekstraksi data umpan balik dari konsumen.

Tabel 2. Hasil ekstraksi data umpan balik

Id responden	Waktu Pengisian Survei	Komentar
12872	12-12-2025	Sudah baik dan memadai.
13037	19-12-2025	Terkadang beberapa CS masih kurang memahami mengenai proses bisnis!!!
14060	19-12-2025	Pendaduan terlalu lambat direspon , dan tidak ada action , hanya sebatas menampung aduan saja
14261	27-11-2025	Kurang puas,,karena masih banyak masalah <i>bug/ error</i> di aplikasi di Playstore yang belum ditindak dan juga aplikasi legal yang sistemnya seperti ilegal

Dari data yang telah terstruktur tersebut kemudian diproses melalui tahap data *cleaning* dan *stemming* untuk meningkatkan kualitas data. Pada tahap *data cleaning & stemming*, karakter yang tidak relevan seperti tanda baca berlebih, simbol, maupun kesalahan penulisan akan dihapus oleh sistem. Kemudian data akan diubah ke bentuk kata dasar sehingga variasi kata yang memiliki makna sama dapat direpresentasikan secara lebih konsisten. Tabel 3 menunjukkan hasil *data cleaning & stemming*.

Tabel 3. Hasil proses *data cleaning & stemming*

Teks Asli	Setelah Cleaning	Setelah Stemming
Sudah baik dan memadai.	sudah baik dan memadai	Sudah baik dan pada
Dari perspektif kami sebagai konsumen, kami menilai pelayanan sudah berjalan efektif dalam memastikan kepatuhan terhadap SOP dan mendukung penerapan tata kelola serta membantu mengenai kendala yang dihadapi konsumen di lapangan.	dari perspektif kami sebagai konsumen kami menilai pelayanan terhadap kualitas layanan telah berjalan efektif dalam memastikan kepatuhan terhadap sop dan mendukung penerapan tata kelola serta membantu menangani kendala yang dihadapi konsumen di lapangan	dari perspektif kami sebagai konsumen kami nilai layanan hadap kualitas layanan telah jalan efektif dalam pasti patuh hadap sop dan dukung terapan tata kelola serta bantu tangani kendala yang hadap konsumen di lapang
Terkadang beberapa cs masih kurang memahami mengenai proses bisnis!!!	terkadang beberapa cs masih kurang memahami mengenai proses bisnis	kadang berapa cs masih kurang paham kena proses bisnis
Pendaduan terlalu lambat direspon , dan tidak ada action , hanya sebatas menampung aduan saja	pendaduan terlalu lambat direspon dan tidak ada action hanya sebatas menampung aduan saja	pengaduan terlalu lambat respon dan tidak ada action hanya sebatas tampung adu saja
Kurang puas,,karena masih banyak masalah <i>bug/ error</i> di aplikasi di Playstore yang belum ditindak dan juga aplikasi legal yang sistemnya seperti ilegal	Kurang puas,,karena masih banyak masalah <i>bug</i> di aplikasi Playstore yang belum ditindak dan juga aplikasi legal yang sistemnya seperti ilegal	Kurang puas,,karena masih banyak masalah <i>bug</i> di aplikasi Playstore yang belum tindak dan aplikasi legal yang sistem seperti ilegal

B. Text Preprocessing

Tahap *text pre-processing* dilakukan untuk menyiapkan data teks agar dapat diproses oleh model analisis sentimen. Proses ini meliputi *tokenization*, *word replacing*, dan *stop word removal*. Tekonisasi dilakukan untuk memecah kalimat menjadi unit kata. Sedangkan *word replacing* dilakukan untuk mengganti kata tidak baku maupun kata asing atau singkatan dengan kata baku sesuai kaidah bahasa Indonesia. Selanjutnya *stop word removal* dilakukan untuk menghilangkan kata-kata umum yang tidak memiliki kontribusi signifikan terhadap pembentukan label sentimen. Hasil dari *text preprocessing* adalah teks umpan balik yang lebih ringkas, bersih, dan representatif terhadap makna sentimen. Tabel 4 menunjukkan hasil *text pre-processing* dalam penelitian.

Tabel 4. Hasil Pra-Pemrosesan Data

Teks	Tokenization	Word Replacing	Penghapusan Stop Word
Sudah baik dan pada	[sudah][baik][dan][pada]	[sudah][baik][dan][pada]	[baik]
dari perspektif kami sebagai konsumen kami nilai layanan hadap kualitas layanan telah jalan efektif dalam pasti patuh hadap sop dan dukung terapan tata kelola serta bantu tangan kendala yang hadap konsumen di lapang	[dari][perspektif][kami][sebagai][konsumen][kami][nilai][layanan][hadap][kualitas][layanan][telah][jalan][efektif][dalam][pasti][patuh][hadap][sop][dan][dukung][terapan][tata][kelola][serta][bantu][tangan][kendala][yang][hadap][konsumen][di][lapang]	[dari][perspektif][kami][sebagai][konsumen][kami][nilai][layanan][hadap][kualitas][layanan][telah][jalan][efektif][dalam][pasti][patuh][hadap][sop][dan][dukung][terapan][tata][kelola][serta][bantu][tangan][kendala][yang][hadap][konsumen][di][lapang]	[perspektif][sebagai][konsumen][nilai][layanan][hadap][kualitas][layanan][telah][jalan][efektif][patuh][hadap][sop][dukung][terapan][tata][kelola][bantu][tangan][kendala][hadap][konsumen][lapang]
kadang berapa cs masih kurang paham kena proses bisnis	[kadang][berapa][cs][masih][kurang][paham][kena][proses][bisnis]	[kadang][berapa][cs][masih][kurang][paham][terhadap][proses][bisnis]	[kadang][cs][kurang][paham][proses][bisnis]
adu terlalu lambat respon dan tidak ada action hanya sebatas tampung adu saja	[adu][terlalu][lambat][respon][dan][tidak][ada][action][hanya][sebatas][menampung][adu][saja]	[pengaduan][terlalu][lambat][respon][dan][tidak][ada][tindakan][hanya][sebatas][tampung][adu][saja]	[pengaduan][lambat][respon][tindakan][menampung][adu][an]
kurang puas karena masih banyak bug/eror di aplikasi playstore yang belum tindak dan juga aplikasi legal yang sistem seperti ilegal	[kurang][puas][karena][masih][banyak][bug/eror][di][aplikasi][playstore][yang][belum][tindak][dan][juga][aplikasi][legal][yang][sistem][seperti][ilegal]	[kurang][puas][karena][masih][banyak][bug/eror][di][aplikasi][il egal][kedok][pinjaman][di][google playstore][yang][belum][ditindak][dan][juga][aplikasi][il egal][dengan][sistem][seperti][ilegal]	[kurang][puas][bug/eror][aplikasi][google playstore][tindakan][aplikasi][legal][sistem][ilegal]

Pada tahap ini, juga diimplementasikan model berbasis *transformer* untuk memanfaatkan pustaka *Hugging Face*. *Hugging face transformer* pada penelitian ini berupa *open source* yang menyediakan model *pre-trained* dalam NLP. Penggunaan pustaka *hugging face* dapat menyederhanakan proses *training* sehingga pelatihan bahasa tidak perlu dilakukan dari awal. Gambar 4 menunjukkan penggunaan pustaka *hugging face transformer* dalam kode *Python*.

```
!pip install transformers
!pip install torch
!pip install pandas
!pip install scikit-learn
```

Gambar 4. Kode implementasi *hugging face transformer*

C. Training & Testing

Tahap ini merupakan inti dari penelitian analisis sentimen konsumen. Pada tahap ini, peneliti melakukan pelatihan dan pengujian termasuk optimasi model klasifikasi hasil sentimen. Peneliti menerapkan arsitektur analisis sentimen dengan mekanisme *rule-based adjustment factor* sebagai inti kontribusi. Pada proses *training* peneliti menerapkan dua metode analisis yaitu *machine learning* konvensional TF-IDF dan SVM. Selanjutnya Indo-BERT diterapkan untuk mengekstraksi makna sentimen dalam bahasa Indonesia. Pendekatan ganda menggunakan kedua metode tersebut dilakukan untuk mendapatkan performa model yang lebih optimal sehingga analisis sentimen dapat dilakukan dengan berbasis konteks.

Pada pendekatan pertama, teks sentimen diubah menjadi representasi numerik menggunakan pendekatan TF-IDF. Melalui metode ini, bobot sentimen pada setiap kata dapat diberikan berdasarkan frekuensi kemunculannya dalam dokumen hasil survei kepuasan konsumen dan keseluruhan korpus. Tabel 5 menunjukkan contoh matriks TF-IDF sebagai hasil pembobotan kata dalam dokumen.

Tabel 5. Matriks TF-IDF

Kata	Dokumen 1	Dokumen 2	Dokumen 3
baik	0.42	0	0
lambat	0	0.36	0.21
kurang	0	0.4	0
sasaran	0	0	0.32

Matriks TF-IDF yang dihasilkan kemudian digunakan sebagai input bagi algoritma SVM untuk membangun model klasifikasi sentimen. SVM digunakan karena memiliki kemampuan yang baik dalam memisahkan kelas pada data berdimensi tinggi. Pada penelitian ini, SVM tepat karena akan mengelola 8.423 data survei dengan kelas sentimen yang dibagi menjadi enam kategori. Keenam kategori kelas sentimen tersebut yaitu 1 (sangat tidak setuju), 2 (tidak setuju), 3 (kurang setuju), 4 (cukup setuju), 5 (setuju), dan 6 (sangat setuju).

Apabila proses klasifikasi awal sudah dilakukan, selanjutnya peneliti menerapkan metode Indo-BERT yang telah dilatih pada korpus bahasa Indonesia. Berbeda dengan TF-IDF yang hanya mempertimbangkan frekuensi kata, Indo-BERT mampu memahami konteks kata sehingga menghasilkan representasi semantik yang lebih kaya. Pada tahap ini, teks yang sudah dalam bentuk token sub-kata kemudian dikonversi menjadi konteks kata atau bisa disebut sebagai *embedding vector*. Tabel 6 menunjukkan hasil analisis sentimen dari survei kepuasan konsumen.

Tabel 6. Hasil analisis sentimen kepuasan konsumen

Kata	Nilai Sentimen	Sentimen
Sudah baik dan memadai.	5	Setuju
Dari perspektif kami sebagai konsumen, kami menilai pelayanan sudah berjalan efektif dalam memastikan kepatuhan terhadap SOP dan mendukung penerapan tata kelola serta membantu mengenai kendala yang dihadapi konsumen di lapangan.	4	Cukup Setuju
Terkadang beberapa cs masih kurang memahami mengenai proses bisnis!!!	3.5	Cukup Setuju
Saya rasa tidak terlalu penting, karena fungsi cs yang dilakukan banyak yang kurang tepat sasaran, harus nya cs lebih banyak kearah untuk memajukan perusahaan yang bisa menaungi kami dari sisi bisnis nya, memberikan saran dan masukan. Belum lagi laporan laporan yang diberikan banyak sekali yang dirasa dari segi kegunaan nya tidak berfaedah!	2	Tidak Setuju
Pendaduan terlalu lambat direspon , dan tidak ada action , hanya sebatas menampung aduan saja	2	Tidak Setuju
Kurang puas,,karena masih banyak <i>bug/ error</i> di aplikasi di Playstore yang belum ditindak dan juga aplikasi legal yang sistemnya seperti ilegal	2	Tidak Setuju

D. Evaluasi

Pada tahap evaluasi, mekanisme *adjustment factor* diberlakukan untuk mendapatkan hasil sentimen yang lebih sesuai dengan konteks yang menjadi perhatian khusus bagi para konsumen. Pada penelitian ini, konsumen memberikan umpan balik terhadap suatu layanan yang diterimanya. Pada nilai survei ditemukan konteks kata yang memiliki tingkat nilai yang lebih tinggi daripada kata lain pada umumnya. Sebagai contoh kata “penanganan *fraud* yang lambat” dengan “penanganan aduan yang lambat”, meski kedua teks tersebut memiliki nilai sentimen yang sama-sama negatif, namun dalam pembobotan, kata “*fraud*” memiliki tingkat yang lebih tinggi dibanding kata “aduan” dalam layanan yang diterima. Dengan demikian kata pertama akan memiliki nilai sentimen yang lebih rendah dibandingkan kata kedua. Gambar 5 menunjukkan baris kode yang berperan sebagai *rule-based adjustment factor*.

```

financial_positive = {
  "stabil":1.5,
  "aman":1.5,
  "transparan":1.6,
  "patuh":1.4,
  "akuntabel":1.4,
  "efisien":1.3,
  "efektif":1.3,
  "terkendali":1.2,
  "terpercaya":1.4,
  "profesional":1.2,
  "optimal":1.2,
  "responsif":1.3,
  "cepat":1.2,
  "akurat":1.3,
  "reliable":1.3,
  "tertib":1.1,
  "terintegrasi":1.2,
  "terjamin":1.3,
  "transaksi lancar":1.4,
  "kredibel":1.4,
  "berkualitas":1.3,
  "memuaskan":1.4,
  "baik":1.2,
  "jelas":1.1,
  "terstruktur":1.2
}

financial_negative = {
  "fraud": -2.0,
  "ilegal": -2.0,
  "risiko": -1.5,
  "kerugian": -1.7,
  "pelanggaran": -1.8,
  "lambat": -1.3,
  "error": -1.4,
  "tidak transparan": -1.7,
  "tidak patuh": -1.8,
  "gagal": -1.5,
  "buruk": -1.3,
  "lemah": -1.2,
  "rentan": -1.4,
  "tidak stabil": -1.6,
  "korupsi": -2.0,
  "penipuan": -2.0,
  "tidak efektif": -1.4,
  "tidak efisien": -1.4,
  "tidak jelas": -1.2,
  "lamban": -1.3,
  "tidak akurat": -1.3,
  "overlap": -1.1,
  "terlambat": -1.3,
  "tidak optimal": -1.2,
  "tidak profesional": -1.4,
  "kurang memahami": -0.5
}

```

Gambar 5. Rule-based adjustment factor

Tabel 7 menunjukkan hasil analisis yang didapatkan dari hasil implementasi metode TF-IDF, SVM, dan Indo-BERT yang telah dipadukan dengan mekanisme *adjustment factor*.

Tabel 7. Hasil analisis sentimen

Kata	Nilai Sentimen Tanpa AF	Sentimen Tanpa AF	Nilai Sentimen dengan AF	Sentimen dengan AF
Sudah baik dan memadai .	5	Setuju	6	Sangat Setuju
Dari perspektif kami sebagai konsumen, kami menilai pelayanan sudah berjalan efektif dalam memastikan kepatuhan terhadap SOP dan mendukung penerapan tata kelola serta membantu mengenai kendala yang dihadapi konsumen di lapangan.	4	Cukup Setuju	5	Setuju
Terkadang beberapa cs masih kurang memahami mengenai proses bisnis!!!	3.5	Cukup Setuju	3	Kurang Setuju
Saya rasa tidak terlalu penting, karena fungsi cs yang dilakukan banyak yang kurang tepat sasaran, harus nya cs lebih banyak kearah untuk memajukan perusahaan yang bisa menaungi kami dari sisi bisnis nya, memberikan saran dan masukan. Belum lagi laporan laporan yang diberikan banyak sekali yang dirasa dari segi kegunaan nya tidak berfaedah!	2	Tidak Setuju	2	Tidak Setuju
Pendaduan terlalu lambat direspon , dan tidak ada action , hanya sebatas menampung aduan saja	2	Tidak Setuju	1	Sangat Tidak Setuju
Kurang puas,,karena masih banyak <i>bug/ eror</i> di aplikasi di Playstore yang belum ditindak dan juga aplikasi legal yang sistemnya seperti ilegal	2	Tidak Setuju	1	Sangat Tidak Setuju

Dapat dilihat pada Tabel 5 bahwa *rule-based adjustment factor* menambahkan pendefinisian faktor dalam penyesuaian berbasis domain kualitas pelayanan untuk meningkatkan akurasi klasifikasi sentimen. Pendekatan ini mempertimbangkan bahwa beberapa istilah dalam sektor layanan jasa memiliki signifikansi semantik yang lebih kuat dibandingkan kata umum lainnya. Integrasi pendekatan ini memungkinkan sistem tidak hanya bergantung pada pola statistik, tetapi juga mempertimbangkan pengetahuan domain dalam proses klasifikasi sentimen. Tabel 8 menunjukkan simulasi perhitungan mekanisme *adjustment factor* yang dapat mempengaruhi nilai sentimen dari penerapan metode TF-IDF, SVM, dan Indo-BERT.

Tabel 8. Mekanisme penerapan *Rule-based adjustment factor*

Kata	TF-IDF, SVM, dan Ino-BERT	Adjustmetn Factor	Sentimen Akhir
Sudah baik dan memadai.	5	+1	6
Dari perspektif kami sebagai konsumen, kami menilai pelayanan sudah berjalan efektif dalam memastikan kepatuhan terhadap SOP dan mendukung penerapan tata kelola serta membantu mengenai kendala yang dihadapi konsumen di lapangan.	4	+1.3	5
Terkadang beberapa cs masih kurang memahami mengenai proses bisnis!!!	3.5	-0.5	3
Saya rasa tidak terlalu penting, karena fungsi cs yang dilakukan banyak yang kurang tepat sasaran, harus nya cs lebih banyak kearah untuk memajukan perusahaan yang bisa menaungi kami dari sisi bisnis nya, memberikan saran dan masukan. Belum lagi laporan laporan yang diberikan banyak sekali yang dirasa dari segi kegunaan nya tidak berfaedah!	2	0	2
Pendaduan terlalu lambat direspon , dan tidak ada action , hanya sebatas menampung aduan saja	2	-1.3	1
Kurang puas,,karena masih banyak bug/ eror aplikasi di Playstore yang belum ditindak dan juga aplikasi legal yang sistemnya seperti Ilegal	2	-2	1

Dalam penelitian ini, penerapan *rule-based adjustment factor* diperluas dengan memasukkan frasa kontekstual yang sering muncul dalam kinerja layanan jasa. Setiap frasa kalimat dapat menggambarkan keunggulan dan keterbatasan terkait proses bisnis, regulasi, tata kelola, hingga pada mekanisme penerapan kualitas layanan jasa. Dengan demikian, integrasi *rule-based adjustment factor* dengan Indo-BERT mampu meningkatkan akurasi klasifikasi sentimen pada dataset survei kepuasan konsumen. Penelitian ini berhasil memberikan kontribusi yang signifikan dalam meningkatkan sensitivitas model terhadap kata-kata kritis dalam jasa manufaktur seperti kualitas layanan, risiko penanganan komplain, dan transparansi penanganan komplain konsumen. Dengan pendekatan ini, sistem mampu menangkap suasana sentimen dengan lebih spesifik terhadap konteks tata kelola jasa dan kualitas layanan perusahaan. Gambar 6 menunjukkan hasil analisis sentimen dari 8.423 data survei konsumen yang memadukan metode TF-IDF, SVM, Indo-BERT, dan *rule-based adjustment factor*.



Gambar 6. Hasil analisis sentimen kepuasan konsumen

Gambar 6 menunjukkan hasil analisis sentimen yang dilakukan terhadap 8.423 data konsumen menyatakan bahwa 13.6% sudah merasa sangat setuju atau sangat puas dengan kualitas pelayanan jasa di perusahaan X. Sedangkan 62,2% merasa setuju, 20,2% cukup setuju, 2,91% kurang setuju, 0,99% tidak setuju, dan 0,05% sangat tidak setuju terhadap layanan yang diterima. Hasil tersebut menunjukkan bahwa pelayanan dalam kualitas layanan jasa sudah baik dan konsumen sudah merasa puas dengan layanan yang diberikan. Jika dilihat

dari hasil analisis sentimen, maka 96% konsumen sudah puas dengan kualitas layanan jasa yang diterima, namun masih ada 4% konsumen merasa tidak puas dengan kualitas layanan jasa yang diterimanya. Nilai tersebut menunjukkan masih terdapat 337 konsumen masih mengharapkan adanya peningkatan kualitas layanan yang dikelola oleh perusahaan.

Dalam mengukur performa sistem analisis sentimen, hasil evaluasi sistem menunjukkan bahwa nilai akurasi klasifikasi dan analisis sentimen sebesar 90%. Nilai tersebut menunjukkan bahwa perpaduan metode TF-IDF, SVM, Indo-BERT, dan *rule-based adjustment factor* memberikan hasil klasifikasi sentimen yang memuaskan. Gambar 7 menunjukkan hasil evaluasi performa metode TF-IDF, SVM, Indo-BERT dan *rule-based adjustment factor* dalam menganalisis sentimen kepuasan konsumen.

```

... Hasil Evaluasi Model Analisis Sentimen
-----
Accuracy : 0.9
Precision : 0.8571428571428571
Recall : 1.0
F1-Score : 0.9230769230769231

Classification Report:

```

	precision	recall	f1-score	support
Negatif	1.00	0.75	0.86	4
Positif	0.86	1.00	0.92	6
accuracy			0.90	10
macro avg	0.93	0.88	0.89	10
weighted avg	0.91	0.90	0.90	10

Gambar 7. Hasil Pengujian Performa Sistem

Berdasarkan hasil pengujian performa sistem, analisis sentimen dengan Indo-BERT menunjukkan performa yang baik hingga mendapatkan nilai akurasi mencapai 90%, presisi 85.7%, *recall* 100%, dan F1-Score 92.3%. Nilai *recall* yang mencapai 100% menunjukkan bahwa seluruh data sentimen pada tahap pengujian berhasil dikenali oleh sistem tanpa adanya sentimen relevan yang terlewat. Hal ini menandakan bahwa integrasi aturan berbasis konteks mampu meningkatkan sensitivitas model dalam mendeteksi berbagai bentuk opini konsumen. Selain itu, nilai F1-Score sebesar 92.3% menunjukkan keseimbangan yang sangat baik antara nilai presisi dan *recall*. Meskipun tingkat presisi berada pada angka 85.7%, sistem tetap mampu mempertahankan tingkat klasifikasi yang stabil dan konsisten terhadap berbagai variasi teks opini. Dengan demikian, kombinasi Indo-BERT dan *rule-based adjustment factor* tidak hanya meningkatkan kemampuan klasifikasi secara kuantitatif, tetapi juga memperkuat interpretasi kontekstual sentimen sehingga lebih relevan digunakan dalam pengukuran kepuasan konsumen dan pengambilan keputusan berbasis data. Peneliti juga membandingkan performa sistem dengan menggunakan metode lain. Tabel 9 menunjukkan evaluasi performa model analisis sentimen.

Tabel 9. Evaluasi Model Analisis Sentimen

Metode	Akurasi	Presisi	Recall	F1-Score
TF-IDF + SVM	84.3%	80.1%	78.6%	79.3%
Indo-BERT	88.7%	84.2%	86.5%	85.3%
TF-IDF + SVM + Indo-BERT	89.1%	84.9%	91.2%	87.9%
TF-IDF + SVM + Indo-BERT + Rule-Based Adjustment Factor	90.0%	85.7%	100%	92.3%

Berdasarkan hasil evaluasi model analisis sentimen kepuasan konsumen diketahui bahwa perpaduan TF-IDF, SVM, Indo-BERT dan *rule-based adjustment factor* menunjukkan peningkatan performa yang signifikan. Perpaduan metode ini mampu memberikan nilai akurasi, presisi, *recall*, dan F1-Score tertinggi jika dibandingkan dengan metode lainnya. Penerapan *rule based-adjustment factor* terbukti mampu meningkatkan sensitivitas model terhadap konteks domain dan intensitas opini konsumen. Pendekatan ini tidak hanya mengandalkan probabilitas prediksi dari model AI, tetapi juga mempertimbangkan bobot kata kunci tertentu yang memiliki pengaruh besar terhadap interpretasi sentimen. Dengan demikian, integrasi metode Indo-BERT, TF-IDF, SVM dan *rule-based adjustment factor* menghasilkan sistem analisis sentimen yang lebih adaptif, kontekstual, dan relevan untuk mendukung pengambilan keputusan organisasi berbasis data.

II. Kesimpulan dan saran

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa perpaduan metode konvensional TF-IDF, SVM, dan Indo-BERT dengan metode kontekstual *adjustment factor* mampu mengidentifikasi pola sentimen pada data konsumen terkait kualitas layanan jasa. Representasi kontekstual yang didapatkan menjadi semakin kuat dan konsisten melalui *rule-based adjustment factor* berdasarkan kata kritis dengan sensitivitas yang lebih tinggi dalam kualitas layanan. Penelitian menunjukkan bahwa 96% konsumen sudah merasa puas dengan kualitas layanan jasa dari perusahaan, namun 4% konsumen masih tidak puas dan mengharapkan adanya peningkatan kualitas layanan jasa. Dalam penggunaan metode ganda dalam penelitian ini juga menunjukkan performa yang memuaskan. Perpaduan metode mampu menghasilkan tingkat akurasi analisis sentimen hingga 90% dan tingkat presisi mencapai 85.7%. Nilai tersebut menunjukkan bahwa perpaduan metode TF-IDF, SVM, Indo-BERT dan *rule-based adjustment factor* dapat memberikan performa analisis sentimen yang sangat baik. Konsistensi penilaian juga menjadi lebih baik karena analisis sentimen juga didasarkan pada kontekstual frasa yang memiliki tingkat sensitivitas lebih kuat dalam konteks kualitas layanan jasa.

Ucapan Terima Kasih

Terimakasih kepada UNS yang telah membantu peneliti dalam memperoleh data dalam penelitian ini.

Daftar Pustaka

- [1] N. L. Ratniasih and N. W. Ninik Jayanti, "Sentiment Analysis of Stakeholder Satisfaction Measurement," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 16, no. 2, p. 217, Apr. 2022, doi: 10.22146/ijccs.72245.
- [2] D. Nuralasari, D. H. Qudsi, N. Chairani, and H. R. Yuliantoro, "Discovering User Sentiment Patterns in Libraries with a Hybrid Machine Learning and Lexicon-Based Approach," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 3, pp. 311–317, Nov. 2024, doi: 10.32736/sisfokom.v13i3.2217.
- [3] Y. Mao, Q. Liu, and Y. Zhang, "Sentiment analysis methods, applications, and challenges: A systematic literature review," Apr. 01, 2024, *King Saud bin Abdulaziz University*. doi: 10.1016/j.jksuci.2024.102048.
- [4] S. J. Pipin and H. Kurniawan, "Analisis Sentimen Kebijakan MBKM Berdasarkan Opini Masyarakat di Twitter Menggunakan LSTM," *SIFO Mikroskil*, vol. 23, no. 2, pp. 197–208, Oct. 2022.
- [5] A. S. Talaat, "Sentiment analysis classification system using hybrid BERT models," *J. Big Data*, vol. 10, no. 1, pp. 1–18, Dec. 2023, doi: 10.1186/s40537-023-00781-w.
- [6] M. Amien and G. F. Gunawan, "ELANG: Journal of Interdisciplinary Research BERT dan Bahasa Indonesia: Studi tentang Efektivitas Model NLP Berbasis Transformer," *ELANG: Journal of Interdisciplinary Research*, vol. 1, no. 2, Jan. 2024, doi: <https://doi.org/10.32664/elang.v1i02>.
- [7] D. I. Putri, A. N. Alfian, M. Y. Putra, and P. D. Mulyo, "IndoBERT Model Analysis: Twitter Sentiments on Indonesia's 2024 Presidential Election," 2024. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [8] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, "Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN," *ILKOM Jurnal Ilmiah*, vol. 14, no. 3, pp. 348–354, Dec. 2022, doi: 10.33096/ilkom.v14i3.1505.348-354.
- [9] Y. A. Singgalen, "Performance Analysis of IndoBERT for Sentiment Classification in Indonesian Hotel Review Data," *Journal of Information System Research (JOSH)*, vol. 6, no. 2, pp. 976–986, Jan. 2025, doi: 10.47065/josh.v6i2.6505.
- [10] Y. Badulescu, F. Cañas, and N. Cheikhrouhou, "Judgmental adjustment of demand forecasting models using social media data and sentiment analysis within industry 5.0 ecosystems," *International Journal of Information Management Data Insights*, vol. 4, no. 2, Nov. 2024, doi: 10.1016/j.jjime.2024.100272.
- [11] C. Sun, J. Mai, W. He, H. Zhu, and Q. Liu, "Complex system modeling using deviation-smoothing belief rule base with training and optimization," *Eng. Appl. Artif. Intell.*, vol. 158, Oct. 2025, doi: 10.1016/j.engappai.2025.111539.
- [12] Y. Yu, B. Li, L. Zhao, C. Ma, and G. Jiao, "A novel optimal control strategy for active suspension systems with control weighting coefficient dynamically adjusted based on operating conditions," *Mech. Syst. Signal Process.*, vol. 244, Jan. 2026, doi: 10.1016/j.ymssp.2025.113752.
- [13] C. Apriyadi and Styawati, "Sentiment Analysis of Cyber Attacks in Bank Syariah Indonesia Using SVM and Indobert Method," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 2, pp. 819–838, Apr. 2025, doi: 10.52436/1.jutif.2025.6.2.2636.
- [14] R. I. Perwira, V. A. Permadi, D. I. Purnamasari, and R. P. Agusdin, "Domain-Specific Fine-Tuning of IndoBERT for Aspect-Based Sentiment Analysis in Indonesian Travel User-Generated Content,"

- Journal of Information Systems Engineering and Business Intelligence*, vol. 11, no. 1, pp. 30–40, Feb. 2025, doi: 10.20473/jisebi.11.1.30-40.
- [15] L. P. F. Larasanty, M. K. A. Jaya, K. W. Astuti, and I. W. M. Santika, “Pengembangan Kuisioner Udayana untuk Penilaian Kepuasan Pasien terhadap Pelayanan Kefarmasian oleh Apoteker di Pusat Kesehatan Masyarakat,” *JURNAL MANAJEMEN DAN PELAYANAN FARMASI (Journal of Management and Pharmacy Practice)*, vol. 11, no. 1, p. 62, Mar. 2021, doi: 10.22146/jmpf.60998.
- [16] L. Murphy, “The Questionnaire Surveying Research Method: Pros, Cons and Best Practices,” *Science Open*, vol. 2, Mar. 2023, doi: 10.14293/S2199-1006.1.SOR-.PP3WYS8.v1.
- [17] M. Siino, I. Tinnirello, and M. La Cascia, “Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers,” *Inf. Syst.*, vol. 121, Mar. 2024, doi: 10.1016/j.is.2023.102342.
- [18] H. T. Kesgin and M. F. Amasyali, “Advancing NLP models with strategic text augmentation: A comprehensive study of augmentation methods and curriculum strategies,” *Natural Language Processing Journal*, vol. 7, Jun. 2024, doi: 10.1016/j.nlp.2024.100071.
- [19] H. H. Hamid, S. Bangare, A. Nirmal, and S. Bangal, “Sentiment Analysis: A Comprehensive Review of Technique and Applications,” *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 10, no. 8, Aug. 2023, [Online]. Available: www.jetir.org
- [20] K. P. Harmandini and K. Muslim, “Analysis of TF-IDF and TF-RF Feature Extraction on Product Review Sentiment,” *Jurnal dan Penelitian Teknik Informatika*, vol. 8, no. 2, 2024, doi: 10.33395/v8i2.13376.
- [21] A. Simanjuntak *et al.*, “Research and Analysis of IndoBERT Hyperparameter Tuning in Fake News Detection,” *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 13, no. 1, pp. 60–67, Feb. 2024, doi: 10.22146/jnteti.v13i1.8532.
- [22] K. Hasanah, “Comparison of Sentiment Analysis Model for Shopee Comments on Google Play Store,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 1, pp. 21–30, Feb. 2024, doi: 10.32736/sisfokom.v13i1.1916.