

**PENERAPAN IndoBERT UNTUK ANALISIS SENTIMEN DAN *LATENT DIRICHLET ALLOCATION* (LDA) UNTUK PEMODELAN TOPIK PADA
ULASAN APLIKASI BANK SUMSELBABEL MOBILE**

SKRIPSI

Sebagai salah satu syarat untuk menyelesaikan studi
di Program Studi Sistem Informasi S1



Oleh

KANASTAIN

09031282126039

**JURUSAN SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA**

2024

HALAMAN PENGESAHAN

SKRIPSI

**PENERAPAN IndoBERT UNTUK ANALISIS SENTIMEN DAN *LATENT*
DIRICHLET ALLOCATION (LDA) UNTUK PEMODELAN TOPIK PADA
ULASAN APLIKASI BANK SUMSELBABEL MOBILE**

Sebagai salah satu syarat untuk penyelesaian

Studi di Program Studi Sistem Informasi S1



Oleh

KANASTAIN 09031282126039

Palembang, 30 Desember 2024

Mengetahui,

Ketua Jurusan Sistem Informasi,



Ahmad Rifai, S.T., M.T.
NIP.197910202010121003

Pembimbing,


Mgs. Afriyan Firdaus, S.Si., M.IT.
NIP.198202122006041003

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Kanastain

NIM : 09031282126039

Program Studi : Sistem Informasi

Judul Skripsi : Penerapan IndoBERT Untuk Analisis Sentimen dan Latent Dirichlet Allocation (LDA) Untuk Pemodelan Topik pada Ulasan Aplikasi Bank Sumselbabel Mobile

Hasil Pengecekan Software iThenticate/Turnitin: 4%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya saya sendiri bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikianlah pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan dari pihak manapun.



Palembang, 30 Desember 2024



Kanastain
Nim. 09031282126039

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada:

Hari : Senin

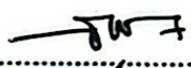



Tanggal : 30 Desember 2024

Nama : Kanastain

NIM : 09031282126039



Judul : Penerapan IndoBERT Untuk Analisis Sentimen dan Latent Dirichlet Allocation (LDA) Untuk Pemodelan Topik pada Ulasan Aplikasi Bank Sumselbabel Mobile

Komisi Penguji:

1. Ketua : Dwi Rosa Indah, M.T. : 
2. Sekretaris : M.Husni Syahbani, M.T. : 
3. Pembimbing : Mgs. Afriyan Firdaus, S.Si., M.IT. : 
4. Penguji : Dedy Kurniawan, M.Sc. : 

Mengetahui,

Ketua Jurusan Sistem Informasi,

Ahmad Rifai, S.T., M.T.
NIP. 197910202010121003

HALAMAN PERSEMBAHAN

MOTTO

"Tidak ada balasan kebaikan kecuali kebaikan (pula)."

(Q.S. Ar-Rahman: 60)

"Do what you can, with what you have, where you are."

(Theodore Roosevelt)

Skripsi ini dipersembahkan untuk:

- ❖ Allah SWT**
- ❖ Orang tua dan keluarga**
- ❖ Seluruh teman dan sahabat seperjuangan angkatan 2021**
- ❖ Dosen pengajar jurusan sistem informasi**
- ❖ Almamater, Universitas Sriwijaya**

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, karena atas rahmat dan karunia-Nya, penulis dapat menyelesaikan skripsi dengan judul " Penerapan IndoBERT Untuk Analisis Sentimen dan Latent Dirichlet Allocation (LDA) Untuk Pemodelan Topik pada Ulasan Aplikasi Bank Sumselbabel *Mobile*". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana Strata Satu (S1) pada Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.

Penulis menyadari bahwa dalam proses penyusunan skripsi ini, banyak pihak yang telah memberikan dukungan dan bantuan, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Allah SWT yang selalu senantiasa memberikan Rahmat dan nikmat berupa ilmu, Kesehatan, dan kesempatan sehingga penulis dapat menyelesaikan tugas akhir ini.
2. Orang tua tercinta yaitu Bapak Sultan Zahril dan Ibu Siptianah, terima kasih sudah berjuang untuk kehidupan penulis, atas setiap keringat yang diteteskan dan setiap doa yang dilangitkan untuk penulis, sehingga dapat melalui proses kehidupan, termasuk dalam menyelesaikan skripsi ini. Terima Kasih atas segala pengorbanan, cinta, dan dukungan yang tida henti diberikan. Semoga Allah SWT selalu memberikan kesehatan kepada kedua orang tua penulis.
3. Bapak Prof.Dr.Erwin, S.Si., M.Si. selaku dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Ahmad Rifai, S.T., M.T. selaku Ketua Jurusan Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya

5. Ibu Iin Seprina, S.KOM., M.KOM. selaku dosen pembimbing Akademik penulis yang telah membimbing, memotivasi, dan memberikan arahan terkait akademik kepada penulis.
6. Bapak Mgs. Afriyan Firdaus, S.Si., M.IT. selaku dosen pembimbing tugas akhir yang telah memebrikan banyak saran dan masukan, serta memberikan bimbingan terbaik sehingga penulis bisa menyelesaikan tugas akhir ini.
7. Ibu Dwi Rosa Indah, M.T. selaku dosen sistem infromasi yang banyak membantu penulis, memberikan masukan serta saran sehingga penulis dapat menyelesaikan tugas akhir.
8. Kepada seseorang yang tak kalah penting, Nabilah Zahirah, terima kasih telah menjadi bagian dari perjalanan penulis yang selalu memberikan dukungan, menghibur, mendengarkan keluh kesah, dan memberikan semangat untuk penulis.
9. Teman seperjuangan penulis dari kelas SIREG C yang telah berjuang dan memiliki tekad yang sama untuk menyelesaikan perkuliahan.
10. The Punggawa, Nopek, Dhani, Imam, Jeki, Fayad, Rizky, robi, amar, aji, dan idris yang telah menjadi bagian terhebat dalam kisah penulis selama masa perkuliahan. Semoga diberikan kelancaran atas segala urusan dan dapat bertemu lagi di lain kesempatan.
11. Himsi sebagai organisasi tempat penulis banyak belajar dan berkembang.
12. Terakhir, terima kasih kepada diri sendiri karena telah mampu menyelesaikan tanggung jawab, berjuang sejauh ini dan berusaha untuk tidak pernah menyerah sampai akhirnya bisa menyelesaikan tugas akhir ini dengan baik.

Penulis menyadari bahwa dalam pembuatan Tugas Akhir ini masih sangat jauh dari kata sempurna. Apabila nantinya ditemukan kekeliruan dalam penelitian Tugas Akhir ini, penulis sangat mengharapkan kritik dan sarannya sebagai bentuk penyempurnaan penelitian ini. Demikian, dengan segala keterbatasan, peneliti mengharapkan Tugas Akhir ini dapat bermanfaat dalam pengetahuan dan wawasan kepada banyak orang, Aamiin Yaa Rabbal Alamin. Semoga semesta selalu mengiringi hal-hal baik dan Langkah kita semua.

Palembang, 30 Desember 2024

Penulis,



Kanastain

09031282126039

**PENERAPAN IndoBERT UNTUK ANALISIS SENTIMEN DAN *LATENT DIRICHLET ALLOCATION* (LDA) UNTUK PEMODELAN TOPIK PADA
ULASAN APLIKASI BANK SUMSELBABEL *MOBILE***

Oleh:

KANASTAIN

09031282126039

ABSTRAK

Mobile banking merupakan salah satu contoh bagaimana layanan perbankan tradisional telah bertransformasi menjadi layanan perbankan digital berkat kemajuan teknologi digital. Aplikasi Bank SumselBabel *Mobile* menjadi fokus utama penelitian untuk memahami persepsi pengguna. Penelitian ini bertujuan untuk menganalisis sentimen menggunakan metode IndoBERT dan BERT sebagai pembanding kinerja, serta mengidentifikasi topik utama dalam ulasan menggunakan *Latent Dirichlet Allocation* (LDA). Data ulasan dikumpulkan pada tanggal 30 Oktober 2024 dengan total 2632 ulasan. IndoBERT memiliki performa yang lebih unggul dibandingkan BERT dalam menganalisis sentimen, dengan *F1-score* 93% berbanding 88%. Hal ini menunjukkan kemampuan IndoBERT lebih baik dalam memahami nuansa bahasa Indonesia dibandingkan BERT. Selanjutnya, pemodelan topik menggunakan LDA pada setiap sentimen menghasilkan nilai *coherence* 0,4559 untuk sentimen negatif, 0,4058 untuk netral, dan 0,4114 untuk sentimen positif.

Kata Kunci : Analisis Sentimen, *mobile banking*, IndoBERT, *Latent Dirichlet Allocation* (LDA), pemodelan topik, Bank Sumsel Babel.

**APPLICATION OF IndoBERT FOR SENTIMENT ANALYSIS AND
LATENT DIRICHLET ALLOCATION (LDA) FOR TOPIC MODELING
ON BANK SUMSELBABEL MOBILE APP REVIEWS**

Oleh:

KANASTAIN (09031282126039)

ABSTRACT

Mobile banking is one example of how traditional banking services have transformed into digital banking services thanks to advances in digital technology. Bank SumselBabel Mobile application is the main focus of research to understand user perception. This research aims to analyze sentiment using the IndoBERT and BERT methods as a performance comparison, as well as identify the main topics in the reviews using Latent Dirichlet Allocation (LDA). The review data was collected on October 30, 2024 with a total of 2632 reviews. IndoBERT has superior performance compared to BERT in analyzing sentiment, with an F1-score of 93% versus 88%. This shows IndoBERT's ability to understand the nuances of Indonesian better than BERT. Furthermore, topic modeling using LDA on each sentiment produces a coherence value of 0.4559 for negative sentiment, 0.4058 for neutral, and 0.4114 for positive sentiment.

Keywords: *Sentiment Analysis, mobile banking, IndoBERT, Latent Dirichlet Allocation (LDA), topic modeling, Bank Sumsel Babel.*

DAFTAR ISI

HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN BEBAS PLAGIAT	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian.....	3
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	5
2.1 Landasan Teori	5
2.1.1 Bank SumselBabel <i>Mobile</i>	5
2.1.2 Analisis Sentimen	5
2.1.3 Pemodelan Topik	6
2.1.4 <i>Google Play Store</i>	7
2.1.5 IndoBERT	7
2.1.6 <i>Latent Dirichlet Allocation (LDA)</i>	9
2.1.7 <i>Web Scraping</i>	10
2.1.8 <i>Preprocessing</i>	11
2.1.11 <i>Lexicon Based</i>	13
2.1.14 <i>Confusion Matrix</i>	13
2.2 Penelitian Terdahulu.....	15
BAB III METODOLOGI PENELITIAN	17
3.1 Objek Penelitian	17

3.2	Jenis Penelitian	17
3.3	Jenis Data	17
3.4	Alur Penelitian.....	17
3.5	Pengumpulan Data	18
3.6	Pra-Pemrosesan Data.....	19
3.6.1	<i>Cleaning</i>	19
3.6.2	<i>Tokenizing</i>	20
3.6.3	Normalization	20
3.6.4	<i>Stopword Removal</i>	21
3.6.5	Stemming	22
3.7	Pelabelan data.....	23
3.8	Pemodelan	23
3.8.1	IndoBERT	24
3.8.2	<i>Latent Dirichlet Allocation (LDA)</i>	27
3.9	Evaluasi	30
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Lingkup Pengembangan	31
4.3	Pra-Pemrosesan Data.....	33
4.3.1	<i>Cleaning</i>	34
4.3.2	<i>Tokenizing</i>	34
4.3.3	Normalization	35
4.3.4	<i>Stop Word Removal</i>	37
4.3.5	<i>Stemming</i>	38
4.4	Pelabelan Data.....	38
4.5	Pemodelan dan Evaluasi Model	39
4.5.1	Analisis Sentimen IndoBERT.....	39
4.5.2	Analisis Sentimen BERT	44
4.5.3	Latent Dirichlet Allocation (LDA)	46
4.6	Hasil dan Analisis.....	49
4.6.1	Evaluasi Model Sentimen	49
4.6.2	<i>Confusion Matrix</i> Model Sentimen.....	52
4.6.3	Rata-Rata <i>F1-Score</i> Model Sentimen	54

4.6.4 Evaluasi Pemodelan Topik	54
4.7 Visualisasi dan Interpretasi Data	58
4.7.1 Analisis Sentimen	58
4.7.2 Pemodelan Topik	60
BAB V KESIMPULAN	64
5.1 Kesimpulan.....	64
5.2 Saran.....	65
DAFTAR PUSTAKA	67

DAFTAR GAMBAR

Gambar 2.1 Arsitektur Dasar IndoBERT	8
Gambar 2.2 Rumus Gibbs sampling.....	9
Gambar 3.1 Alur Penelitian	17
Gambar 3.2 Contoh Ulasan Positif.....	18
Gambar 3.3 Contoh Ulasan Negatif	19
Gambar 4.1 Sampel Ulasan BSB <i>Mobile</i>	33
Gambar 4.2 <i>F1-score</i> Berdasarkan Sentimen.....	51
Gambar 4.3 <i>Confusion Matrix</i> IndoBERT	53
Gambar 4.4 <i>Confusion Matrix</i> BERT.....	53
Gambar 4.5 Rerata <i>F1-score</i> Tiap Model.....	54
Gambar 4.6 Persentase Keseluruhan Sentimen	59
Gambar 4.7 Visualisasi pyLDAvis Sentimen Negatif.....	60
Gambar 4.8 Visualisasi pyLDAvis Sentimen Netral.....	61
Gambar 4.9 Visualisasi pyLDAvis Sentimen Positif	62

DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i>	14
Tabel 2.2 Penelitian Terdahulu.....	15
Tabel 3.1 Contoh <i>Cleaning</i>	19
Tabel 3.2 Contoh <i>Tokenizing</i>	20
Tabel 3.3 Contoh <i>Normalization</i>	21
Tabel 3.4 Contoh <i>Stopword Removal</i>	22
Tabel 3.5 Contoh <i>Stemming</i>	23
Tabel 3.6 Contoh <i>TOKENizing</i> IndoBERT	26
Tabel 4.1 Pustaka Python	31
Tabel 4.2 Ulasan Hasil <i>Cleaning</i>	34
Tabel 4.3 Ulasan Hasil Tokenizing	35
Tabel 4.4 Hasil <i>Normalization</i>	37
Tabel 4.5 Hasil <i>Normalization</i>	37
Tabel 4.6 Hasil <i>Stemming</i>	38
Tabel 4.7 Distribusi Label Sentimen	39
Tabel 4.8 Hasil Evaluasi Model IndoBERT.....	50
Tabel 4.9 Hasil Evaluasi Model BERT	50
Tabel 4.10 <i>Coherence Score</i> LDA	55
Tabel 4.11 Topik Utama Setiap Sentimen.....	56

DAFTAR LAMPIRAN

Lampiran 1 Surat Pengantar Penelitian	A-1
Lampiran 2 Sampel <i>Dataset</i> Hasil Pelabelan	B-1
Lampiran 3 Perbandingan Rating Aplikasi	C-1
Lampiran 4 Form Wawancara Pengguna	D-1
Lampiran 5 Turnitin	E-1
Lampiran 6 Kartu Konsultasi	F-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Industri perbankan mengalami transformasi besar dengan kemajuan teknologi informasi, termasuk penerapan *mobile banking* yang mempermudah pengguna dalam berbagai transaksi keuangan, seperti pembayaran tagihan, transfer dana, dan pemantauan rekening (Miftahusalam et al., 2023). Bank Sumsel Babel, sebagai bank pembangunan daerah, telah menghadirkan aplikasi Bank SumselBabel *Mobile* untuk memenuhi kebutuhan layanan perbankan digital yang cepat dan sederhana.

Aplikasi Bank SumselBabel *Mobile* telah diunduh lebih dari 500 ribu kali di Google Play Store, menunjukkan tingginya minat dan kepercayaan pengguna. Namun, meskipun ulasan pengguna di Google Play Store diperhatikan oleh pihak manajemen, proses evaluasinya belum dilakukan secara sistematis. Hal ini berpotensi mengabaikan isu-isu penting yang muncul berulang kali. Beragam ulasan juga mencakup keluhan teknis seperti kesulitan *login* dan apresiasi terhadap fitur-fitur yang membantu transaksi harian. Selain itu, rating aplikasi Bank SumselBabel *Mobile* masih relatif lebih rendah jika dibandingkan dengan aplikasi *mobile banking* dari bank pembangunan daerah lainnya.

Untuk memahami persepsi pengguna secara menyeluruh, diperlukan analisis sentimen untuk mengelompokkan ulasan berdasarkan kategori positif, netral, atau negatif. Analisis sentimen membantu bank dalam mengidentifikasi area perbaikan serta mengetahui tingkat kepuasan pengguna (Larasati et al., 2022).

Selain itu, pemodelan topik menggunakan *Latent Dirichlet Allocation* (LDA) dapat menggali tema utama dalam ulasan pengguna, seperti performa aplikasi, keamanan, atau kemudahan penggunaan (Kustiyahningsih & Permana, 2024).

Untuk menganalisis sentimen ulasan berbahasa Indonesia secara mendalam, diperlukan model yang dirancang khusus untuk bahasa Indonesia. IndoBERT, pengembangan dari BERT-base, terbukti unggul dalam memahami struktur kalimat dan konteks lokal. Penelitian sebelumnya menunjukkan performa IndoBERT yang sangat baik, seperti akurasi 96% pada analisis sentimen aplikasi di bidang kesehatan (Imaduddin et al., 2023) dan akurasi hingga 89,84% pada ulasan e-commerce (Cindy Pradhisa & Fajriyah, 2024). Selain itu, untuk mengidentifikasi tema utama dalam ulasan, metode *Latent Dirichlet Allocation* (LDA) menjadi pilihan yang efektif. LDA mampu menemukan pola hubungan kata dalam dokumen, sehingga dapat mengidentifikasi aspek-aspek yang menjadi perhatian utama pengguna, seperti performa, keamanan, dan fitur aplikasi (Merawati et al., 2021).

Penelitian ini bertujuan untuk menganalisis sentimen dan memodelkan topik dari ulasan pengguna aplikasi Bank SumselBabel *Mobile* di Google Play Store dengan menggunakan metode IndoBERT untuk analisis sentimen dan *Latent Dirichlet Allocation* (LDA) untuk pemodelan topik, serta membandingkan kinerja IndoBERT dengan model BERT. Hasil dari penelitian ini diharapkan dapat memberikan wawasan mendalam mengenai persepsi nasabah terhadap aplikasi, serta mengidentifikasi topik utama yang sering dibahas pengguna, sehingga dapat membantu Bank Sumsel Babel dalam meningkatkan kualitas layanan dan memperbaiki fitur-fitur aplikasi secara tepat sasaran.

Berdasarkan uraian yang telah dijabarkan di atas, maka penulis melakukan penelitian dengan judul “**PENERAPAN IndoBERT UNTUK ANALISIS SENTIMEN DAN *LATENT DIRICHLET ALLOCATION (LDA)* UNTUK PEMODELAN TOPIK PADA ULASAN APLIKASI BANK SUMSELBABEL MOBILE**”.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana analisis sentimen pengguna terhadap aplikasi Bank SumselBabel *Mobile* menggunakan model IndoBERT?
2. Bagaimana perbandingan kinerja model IndoBERT dan BERT dalam analisis sentimen berbahasa Indonesia pada ulasan pengguna aplikasi Bank SumselBabel *Mobile*?
3. Bagaimana pemodelan topik dari ulasan pengguna aplikasi Bank SumselBabel *Mobile* menggunakan model *Latent Dirichlet Allocation (LDA)*?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut :

1. Melakukan analisis sentimen ulasan pengguna aplikasi Bank SumselBabel *Mobile* menggunakan model IndoBERT.
2. Membandingkan kinerja model IndoBERT dan BERT dalam klasifikasi sentimen ulasan pengguna berbahasa Indonesia.

3. Mengidentifikasi topik-topik utama dari ulasan pengguna disetiap sentimen melalui pemodelan topik menggunakan model *Latent Dirichlet Allocation* (LDA).

1.4 Manfaat Penelitian

Manfaat yang dapat diambil dari penelitian ini adalah sebagai berikut :

1. Hasil penelitian dapat dijadikan sebagai salah satu bahan evaluasi untuk meningkatkan kualitas layanan digital pada aplikasi Bank SumselBabel *Mobile*.
2. Penelitian ini bisa digunakan sebagai referensi dan alternatif sumber bagi penelitian sejenis di masa mendatang.

1.5 Batasan Masalah

Untuk menghindari terjadinya pembahasan yang menyimpang dari judul ataupun topik penelitian, maka peneliti membatasi ruang masalah penelitian sebagai berikut:

1. Penelitian ini menggunakan data ulasan pengguna Google Play Store pada aplikasi Bank SumselBabel *Mobile* yang tersedia dari tanggal 29 Oktober 2023 hingga 29 Oktober 2024.
2. Analisis hanya dilakukan terhadap ulasan berbahasa Indonesia.
3. Penelitian ini menggunakan model IndoBERT untuk melakukan analisis sentimen dan model *Latent Dirichlet Allocation* (LDA) untuk pemodelan topik.

DAFTAR PUSTAKA

- Abdul Fatah, D., Ihsan Kamil, F., Soesilo, B., & Sistem Informasi, P. (2024). PENERAPAN NAÏVE BAYES DAN LATENT DIRICHLET ALLOCATION (LDA) UNTUK ANALISIS SENTIMEN DAN PEMODELAN TOPIK PADA PROYEK KERETA CEPAT JAKARTA-BANDUNG APPLICATION OF NAÏVE BAYES AND LATENT DIRICHLET ALLOCATION (LDA) FOR SENTIMENT ANALYSIS AND TOPIC MODELING ON THE JAKARTA-BANDUNG HIGH-SPEED RAIL PROJECT. *Jurnal SimanteC*, 12(2).
- Aida Sapitri, I., Fikry, M., Sains dan Teknologi, F., & Islam Negeri Sultan Syarif Kasim Riau, U. (2023). PENGKLASIFIKASIAN SENTIMEN ULASAN APLIKASI WHATSAPP PADA GOOGLE PLAY STORE MENGGUNAKAN SUPPORT VECTOR MACHINE. *Jurnal TEKINKOM*, 6(1). <https://doi.org/10.37600/tekinkom.v6i1.773>
- Amanda, A., & Nurmawati, E. (2023). ANALISIS SENTIMEN DAN PEMODELAN TOPIK PADA TWEET TERKAIT DATA BADAN PUSAT STATISTIK. *Jurnal Sistem Informasi Dan Informatika (Simika) P-ISSN*, 6(2), 2622–6901.
- Cindy Pradhisa, K., & Fajriyah, R. (2024). Analisis Sentimen Ulasan Pengguna E-commerce di Google Play Store Menggunakan Metode IndoBERT. *Technology and Science (BITS)*, 6(1). <https://doi.org/10.47065/bits.v6i1.5247>
- Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using IndoBERT Language Models. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, 9(3), 746–757. <https://doi.org/10.26555/jiteki.v9i3.26490>
- Imaduddin, H., Yusfida A'la, F., & Nugroho, Y. S. (2023). Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach. *IJACSA) International Journal of Advanced Computer Science and Applications*, 14(8), 2023. www.ijacsa.thesai.org
- Khairunnisa, S., Adiwijaya, A., & Faraby, S. Al. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2), 406. <https://doi.org/10.30865/mib.v5i2.2835>

- Khder, M. A. (2021). Web scraping or web crawling: State of art, techniques, approaches and application. *International Journal of Advances in Soft Computing and Its Applications*, 13(3), 144–168. <https://doi.org/10.15849/ijasca.211128.11>
- Kustiyahningsih, Y., & Permana, Y. (2024). Penggunaan Latent Dirichlet Allocation (LDA) dan Support-Vector Machine (SVM) Untuk Menganalisis Sentimen Berdasarkan Aspek Dalam Ulasan Aplikasi EdLink. *Teknika*, 13(1), 127–136. <https://doi.org/10.34148/teknika.v13i1.746>
- Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(9), 4305–4313. <http://j-ptiik.ub.ac.id>
- Matira, Y., & Setiawan, I. (2023). Pemodelan Topik pada Judul Berita Online Detikcom Menggunakan Latent Dirichlet Allocation. *Estimasi: Journal of Statistics and Its Application*, 4(1), 2721–379. <https://doi.org/10.20956/ejsa.vi.24843>
- Miftahusalam, A., Pratiwi, H., & Slamet, I. (2023). Perbandingan Metode Random Forest dan Naive Bayes pada Analisis Sentimen Review Aplikasi BCA Mobile. *Prosiding Seminar Nasional*.
- Miranda, C. H., Sanchez-Torres, G., & Salcedo, D. (2023). Exploring the Evolution of Sentiment in Spanish Pandemic Tweets: A Data Analysis Based on a Fine-Tuned BERT Architecture. *Data*, 8(6). <https://doi.org/10.3390/data8060096>
- Musfiroh, D., Khaira, U., Eko, P., Utomo, P., & Suratno, T. (2021). Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1, 24–33.
- Mutmainah, S., Fudholi, D. H., & Hidayat, S. (2023). Analisis Sentimen dan Pemodelan Topik Aplikasi Telemedicine Pada Google Play Menggunakan BiLSTM dan LDA. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(1), 312. <https://doi.org/10.30865/mib.v7i1.5486>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 5, Issue 2).

- Ramadhani, B., & Suryono, R. R. (2024). Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(2), 714. <https://doi.org/10.30865/mib.v8i2.7458>
- Rifaldi, D., Abdul Fadlil, & Herman. (2023). Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet “Mental Health.” *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 161–171. <https://doi.org/10.51454/decode.v3i2.131>
- Supriyanto, J., Alita, D., & Isnain, A. R. (2023). Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring. *Jurnal Informatika Dan Rekayasa Perangkat Lunak*, 4(1), 74–80. <https://doi.org/10.33365/jatika.v4i1.2468>
- Syam, Abd. A., Hardy M, G., Salim, A., Suriyanto, D. F., & Fajar B, M. (2024). ANALISIS TEKNIK PREPROCESSING PADA SENTIMEN MASYARAKAT TERKAIT KONFLIK ISRAEL-PALESTINA MENGGUNAKAN SUPPORT VECTOR MACHINE. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9(3), 1464–1472. <https://doi.org/10.29100/jipi.v9i3.5527>
- Vindua, R., & Zailani, A. U. (2023). Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python. *JURIKOM (Jurnal Riset Komputer)*, 10(2), 479. <https://doi.org/10.30865/jurikom.v10i2.5945>
- Zakeshia Tiara Kannitha, D., Mustafid, & Kartikasari, P. (2022). PEMODELAN TOPIK PADA KELUHAN PELANGGAN MENGGUNAKAN ALGORITMA LATENT DIRICHLET ALLOCATION DALAM MEDIA SOSIAL TWITTER. *JURNAL GAUSSIAN*, 11(2), 266–277. <https://ejournal3.undip.ac.id/index.php/gaussian/>