

# Uncovering Hidden Sentiments and Topics in Online Lending Application Reviews with the Valence Aware Dictionary and sEntiment Reasoner (VADER) and Latent Dirichlet Allocation (LDA) Approaches

Fikri Fahru Roji<sup>1\*</sup>, Windi Ariesti Anggraeni<sup>2</sup>, Riyad Sabilul Muminin<sup>3</sup>,  
Dendi Ramdani<sup>4</sup>, Yayan Cahyan<sup>5</sup>

<sup>1</sup>Program Studi Bisnis Digital, Fakultas Ekonomi, Universitas Garut, Garut, Jawa Barat, Indonesia

<sup>2</sup>Program Studi Akuntansi Fakultas Ekonomi, Universitas Garut, Garut, Jawa Barat, Indonesia

<sup>3</sup>Program Studi Rekayasa Perangkat Lunak, Fakultas Komunikasi dan Informasi, Universitas Garut, Garut, Jawa Barat, Indonesia

<sup>4</sup>Program Studi Sistem Informasi, Institut Pendidikan Indonesia, Garut, Jawa Barat, Indonesia

<sup>5</sup>Program Studi Manajemen Informatika, Fakultas Teknologi Informasi, Universitas Sebelas April, Sumedang, Jawa Barat, Indonesia

\*Penulis koresponden, e-mail : [fikri@uniga.ac.id](mailto:fikri@uniga.ac.id)

---

**Abstract:** *Online lending (pinjol) has become an important part of the digital transformation of the financial sector, offering people easy access to funds. However, the increasing reliance on user reviews as a decision-making factor raises concerns about their authenticity and credibility. This research aims to analyze the sentiments and topics that appear in the reviews of Akulaku, Kredivo, and EasyCash lending apps on the Google Play Store. Using text mining techniques, VADER sentiment analysis, and LDA topic modeling, this research reveals dominant positive sentiments related to ease of use, service speed, and customer support. However, there were also negative reviews regarding loan application difficulties, technical issues, and bad experiences with billing and payments. This research provides valuable insights into the preferences and concerns of pinjol users, which can serve as a reference for service providers to improve the quality of their products and services.*

**Keywords:** *Sentiment Analysis; Topic Modeling; Lending Apps; VADER; LDA*

**Abstrak:** Pinjaman online (pinjol) telah menjadi bagian penting dalam transformasi digital sektor keuangan, menawarkan akses mudah ke dana bagi masyarakat. Namun, meningkatnya ketergantungan pada ulasan pengguna sebagai faktor pengambilan keputusan memunculkan kekhawatiran tentang keaslian dan kredibilitas ulasan tersebut. Penelitian ini bertujuan untuk menganalisis sentimen dan topik yang muncul dalam ulasan aplikasi pinjol Akulaku, Kredivo, dan EasyCash di Google Play Store. Dengan menggunakan teknik text mining, analisis sentimen VADER, dan pemodelan topik LDA, penelitian ini mengungkap sentimen positif yang dominan terkait kemudahan penggunaan, kecepatan layanan, dan dukungan pelanggan. Namun, terdapat juga ulasan negatif mengenai kesulitan pengajuan pinjaman, masalah teknis, dan pengalaman buruk terkait penagihan dan pembayaran. Penelitian ini memberikan wawasan berharga tentang preferensi dan masalah pengguna pinjol, yang dapat menjadi acuan bagi penyedia layanan untuk meningkatkan kualitas produk dan layanan mereka.

**Kata kunci:** Analisis Sentimen; Pemodelan Topik; Pinjaman Online; VADER; LDA

---

## PENDAHULUAN

Kemajuan teknologi yang pesat telah mendorong integrasi pinjaman online (pinjol) sebagai komponen penting dalam transformasi digital di sektor keuangan (Carranza et al. 2021;

Ruangkjanases et al. 2021; Ullah et al. 2022). Pinjaman online, yang muncul sebagai solusi keuangan yang inovatif, menawarkan akses yang mudah terhadap dana tanpa memerlukan proses yang rumit, sehingga masyarakat modern semakin cenderung mengadopsi platform pinjaman online sebagai alternatif yang efisien dan mudah diakses (Zhao, Peng, and Li 2022). Lanskap regulasi juga telah berevolusi untuk mengakomodasi peningkatan pinjaman daring, yang menandakan pergeseran ke arah era yang lebih terregulasi di industri ini (Chung et al. 2023; Huang and Pontell 2023). Pemberlakuan langkah-langkah regulasi, seperti Peraturan Otoritas Jasa Keuangan Nomor 10/POJK.05/2022 tentang Layanan Pendanaan Bersama Berbasis Teknologi Informasi, telah memainkan peran penting dalam membentuk lanskap pinjaman daring. Selain itu, kemunculan pinjaman konsumen online dapat dikaitkan dengan faktor-faktor seperti biaya operasional yang rendah, arbitrase regulasi, dan pelanggaran (Xia et al. 2022).

Pertumbuhan pinjaman online diiringi dengan meningkatnya ketergantungan pada ulasan dan testimoni pengguna sebagai faktor pengambilan keputusan utama bagi calon peminjam (Ming Wang Ge Zhan and Meng 2022). Peran ulasan pengguna dalam aplikasi pinjaman daring menjadi semakin signifikan, dengan individu menggunakan ulasan untuk mendapatkan wawasan, mengevaluasi pengalaman pengguna sebelumnya, dan membuat keputusan yang tepat mengenai aplikasi yang mereka pilih untuk digunakan (Barbado, Araque, and Iglesias 2019; Gonzalez 2023; Stein et al. 2020). Namun, kualitas dan keaslian ulasan-ulasan tersebut mulai dipertanyakan, karena beberapa pengguna menggunakan praktik yang tidak etis, seperti menulis ulasan palsu, untuk meningkatkan reputasi online mereka (Vijayarathy and Butler 2016). Fenomena ini menimbulkan kekhawatiran tentang kredibilitas dan keandalan testimoni pengguna dalam proses pengambilan keputusan (Dong et al. 2020; Vijayarathy and Butler 2016).

Untuk mengatasi tantangan ini, penelitian ini mengusulkan penggunaan *Natural Language Processing* (NLP), analisis teks, dan analisis sentimen untuk mengklasifikasikan ulasan pengguna ke dalam berbagai tujuan, seperti pemberian informasi, pencarian informasi, permintaan fitur, dan penemuan masalah (Noei, Zhang, and Zou 2021). Selain itu, penggunaan model topik, seperti *Latent Dirichlet Allocation* (LDA) dapat membantu dalam mengelompokkan kalimat dan mengidentifikasi kategori yang terkait dengan fungsionalitas aplikasi yang diberikan (Okuboyejo and Koyejo 2021; Puspita, Shiddieq, and Roji 2024; Roji et al. 2023). Penggunaan *word embedding*, seperti Word2Vec, dapat meningkatkan kinerja model topik, memungkinkan eksplorasi yang lebih akurat terhadap perilaku permintaan konsumen (Zhuo 2022).

Sifat teks yang tidak terstruktur dalam ulasan pengguna menggarisbawahi perlunya metode lanjutan, seperti model *Neural Network*, untuk menganalisis permintaan pengguna dan menentukan topik utama dalam ulasan (Gribkov and Yekhlakov 2020). Model analisis sentimen dapat mengidentifikasi konten data teks sebagai positif, negatif, atau netral, dan termasuk dalam bidang

NLP (Aditya, Supriyati, and Listyorini 2022). Penggunaan *text mining* dan pendekatan pengambilan keputusan multi kriteria dapat memberikan wawasan yang berharga untuk mengukur kepuasan pengguna terhadap aplikasi (Dina et al. 2021). Analisis ulasan pengguna dalam aplikasi pinjaman online membutuhkan pendekatan yang beragam, dengan menggabungkan teknik-teknik canggih seperti pemrosesan bahasa alami, analisis sentimen, dan pemodelan topik. Metodologi-metodologi ini dapat memberikan wawasan berharga tentang niat pengguna, kepuasan, dan pengalaman pengguna secara keseluruhan, yang pada akhirnya berkontribusi pada pengambilan keputusan yang tepat bagi calon peminjam.

Untuk mengatasi tantangan kompleksitas bahasa manusia, metode analisis sentimen yang dapat diandalkan sangat dibutuhkan. Salah satu solusi yang akan diadopsi dalam penelitian ini adalah metode *Valence Aware Dictionary and sEntiment Reasoner* (VADER), yang dirancang khusus untuk analisis sentimen dalam Bahasa Inggris (Davis, Gillmore, and Millard 2020). VADER mengandalkan leksikon dan lima aturan umum untuk memetakan fitur leksikal ke skor sentimen (Long et al. 2023). VADER telah digunakan untuk menghitung skor gabungan untuk menentukan sentimen dan mengkategorikannya sebagai positif, netral, atau negatif (Anggraeni, Fahru Roji, and Alkautsar 2023; Jung et al. 2023; Lee, Lee, and Kim 2022; Liu and Liu 2021). Melalui penerapan pemodelan topik dengan LDA dan analisis sentimen dengan VADER, penelitian ini diharapkan dapat memberikan wawasan mendalam tentang preferensi pengguna, isu-isu umum, dan kualitas layanan yang dapat meningkatkan pengalaman pengguna pada aplikasi pinjaman online.

## **KAJIAN PUSTAKA**

### **A. Pinjaman Online**

Pinjaman online telah menjadi salah satu inovasi paling menonjol dalam sektor keuangan digital. Dengan kemajuan teknologi informasi dan komunikasi, layanan pinjaman online menjadi lebih mudah diakses dan lebih efisien dibandingkan dengan lembaga keuangan tradisional. Pinjaman online memungkinkan individu untuk mengajukan dan menerima dana dalam waktu singkat tanpa perlu melalui proses administrasi yang panjang dan rumit. Hal ini memberikan solusi keuangan yang praktis bagi banyak orang, terutama mereka yang tidak memiliki akses ke layanan perbankan konvensional. Pertumbuhan pesat platform pinjaman online didorong oleh beberapa faktor utama. Biaya operasional yang lebih rendah memungkinkan penyedia layanan untuk menawarkan bunga yang lebih kompetitif. Proses persetujuan yang cepat dan kemudahan akses melalui aplikasi mobile juga menjadi daya tarik utama bagi pengguna (Carranza et al. 2021). Selain itu, lanskap regulasi yang berkembang untuk mengakomodasi peningkatan pinjaman daring turut berkontribusi pada pertumbuhan sektor ini (Huang and Pontell 2023).

## B. Analisis sentimen

Teknik analisis sentimen digunakan untuk mengevaluasi opini, penilaian, sikap, dan emosi yang diungkapkan oleh individu mengenai suatu entitas, baik itu produk, layanan, organisasi, maupun topik tertentu. (Salsabila, Sihombing, and Sitorus 2022). Metode yang dipakai dalam analisis sentimen bisa berbeda-beda, seperti *Naïve Bayes*, *Support Vector Machine (SVM)*, dan algoritma lain dalam *Machine Learning* (Wibowo, Darmawan, and Amalia 2022). (As-Salafiyah dan Rusydiana 2022), propaganda publik (Sihombing dan Setiawan 2022), kebutuhan fitur fast track pada platform streaming (Waluyan dan Hartomo 2022), opini publik mengenai vaksin COVID-19 (Prabowo, Pramunendar, dan Megantara 2022), dan berbagai topik lainnya. Dengan penerapan teknik analisis sentimen, studi ini mampu menyediakan wawasan serta informasi yang relevan di berbagai bidang dan situasi.

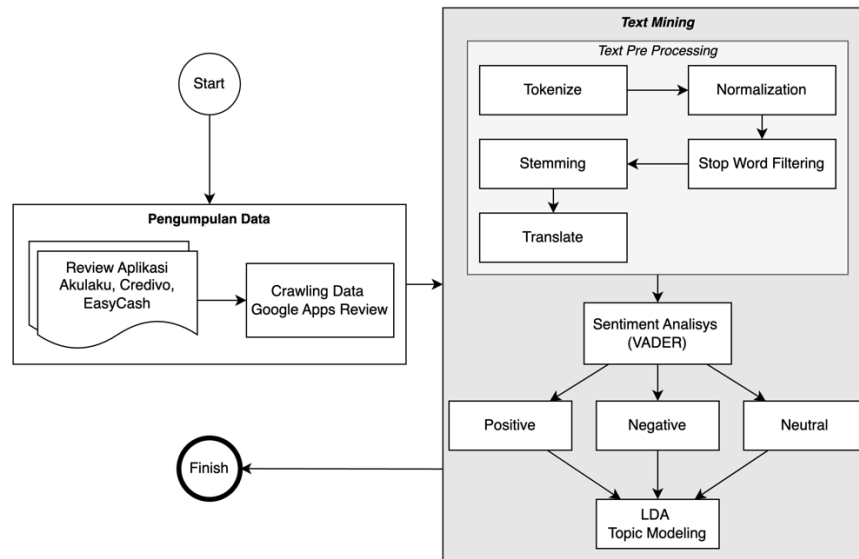
## C. Pemodelan Topik

Teknik pemodelan topik dalam text mining memungkinkan penemuan pola dan hubungan tersembunyi dalam data serta dokumen teks. Teknik ini telah banyak diterapkan di berbagai disiplin ilmu, termasuk rekayasa perangkat lunak, ilmu politik, ilmu kedokteran, dan linguistik. Salah satu teknik populer untuk pemodelan topik adalah *Latent Dirichlet Allocation (LDA)*, yang telah banyak digunakan dan diteliti dalam area ini. (Jelodar et al. 2018). Para peneliti telah mengusulkan model-model berbeda berdasarkan LDA untuk meningkatkan pemodelan topik. Pemodelan topik telah diterapkan untuk merekomendasikan artikel ilmiah, mengelompokkan dokumen, meningkatkan pencarian bibliografis, dan menganalisis volume besar informasi yang tak terstruktur (Zuluaga 2023). Selain itu, pemodelan topik telah digunakan dalam domain khusus seperti literatur hidrologi, ilmu bahan, dan penelitian pemasaran (Masyhuri 2022). Penerapan pemodelan topik telah menunjukkan hasil yang menjanjikan dalam hal kualitas topik, kemampuan interpretasi, dan performa prediksi (Dieng, Vidal, and Blei 2020). Secara keseluruhan, pemodelan topik telah menjadi teknik yang serbaguna dengan berbagai aplikasi dan telah berkontribusi pada kemajuan berbagai bidang penelitian.

## METODE PENELITIAN

Penelitian ini bertujuan untuk menginvestigasi sentimen dan topik yang muncul dari ulasan aplikasi Akulaku, Kredivo, dan EasyCash di Google App Store. Dengan menerapkan teknik *text mining*, penelitian ini diharapkan dapat mengungkap bagaimana pengguna mempersepsikan aplikasi-aplikasi tersebut, mengidentifikasi fitur yang disukai atau tidak disukai, serta masalah yang sering menjadi keluhan. Informasi yang diperoleh sangat penting bagi calon pengguna untuk memutuskan penggunaan aplikasi dan bagi pengembang untuk meningkatkan kualitas produk dan layanan mereka. Tahapan penelitian ini diilustrasikan pada Gambar 1, dimulai dengan pengumpulan data dari ulasan aplikasi, dilanjutkan dengan pra-pemrosesan data untuk membersihkan kata-kata

yang tidak relevan. Langkah berikutnya adalah melakukan analisis sentimen untuk setiap ulasan menggunakan VADER, dan kemudian melakukan pemodelan topik dengan menggunakan *Latent Dirichlet Allocation (LDA)*.



Gambar 1. *Research Methodology*

### A. Pengumpulan Data

Proses pengambilan data ulasan dari Google Play Store dilakukan menggunakan bahasa pemrograman Python dan pustaka Google Play Scraper. Pustaka ini menyediakan API tidak resmi untuk mengakses Google Play Store (Amir Latif et al. 2019), yang memungkinkan penulis untuk mengekstrak dan memanipulasi data ulasan dari aplikasi atau kategori tertentu (Munandar, Farikhin, and Widodo 2023). Pengumpulan data dalam penelitian ini merupakan pengambilan data primer langsung dari sumber asli, yang khusus dikumpulkan oleh peneliti untuk menjawab pertanyaan penelitian. Total terdapat **419.691 ulasan** yang dikumpulkan secara keseluruhan. Rinciannya adalah 128.156 ulasan untuk aplikasi Akulaku, 177.309 ulasan untuk Kredivo, dan 114.226 ulasan untuk EasyCash. Data yang dikumpulkan mencakup sepuluh properti utama: ReviewId, userName, user image, content, at, score, thumbsUpCount, reviewCreateVersion, responsecontent, dan RespondAt, sebagaimana ditunjukkan dalam Gambar 2.

	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreateVersion	at	replyContent	repliedAt	appVersion	source
1	c9bf066f-e	IsiKode-A	<a href="https://play">https://play</a>	Kekinian !!!	4	40	5.0.01	2024-01-19 09.44.53			5.0.01	akulaku
2	c35ef841-	isiRep_	<a href="https://play">https://play</a>	kalau menu	5	49	5.0.01	2024-01-19 02.38.17			5.0.01	akulaku
3	6136ab38-	Rindi Antika	<a href="https://play">https://play</a>	Buat apk ny	4	742	4.9.91	2024-01-06 13.52.35			4.9.91	akulaku
....	....	....	....	....	....	....	....	....	....	....	....	....
419689	1deb44bd-	A Nur Al Huda		Gak pernah	5	0	3.2.8	2020-09-10 16.29.30	Hai, terima k	2020-09-10 17.08.36	3.2.8	easycaash
419690	4a2d8f9c-	Reza Mastur		Perlu di cob	5	0	3.4.8	2021-05-10 00.31.50			3.4.8	easycaash
419691	5df224cf-e	Ruston. 808		semoga bu	3	0	3.7.0	2021-08-12 12.17.49			3.7.0	easycaash

Gambar 2. Hasil Proses Pengumpulan Data

## B. *Text Preprocessing*

Pada tahap *Preprocessing* penelitian ini dilakukan menggunakan Bahasa pemrograman Python dan Microsoft Excel sebagai tempat penyimpanan data. Proses *preprocessing* data ini krusial untuk memudahkan analisis sentimen dan meningkatkan efektivitas serta efisiensi dalam pengolahan data (Nofiyani and Wulandari 2022). Tahapan proses data *preprocessing* pada sentiment analisis melibatkan beberapa langkah untuk mempersiapkan data sebelum dilakukan analisis sentimen. Beberapa tahapan yang umum dilakukan dalam preprocessing data adalah *case folding*, *tokenizing*, *normalization*, *filtering (stopword removal)*, dan *stemming* (Utami and Artana 2022). Data Transformasi dengan tahapan *Case Folding*, yaitu data-data tersebut akan diubah menjadi sama (lowercase). *Data Cleaning* dengan tahapan *Filtering*, yaitu mengeliminasi kata-kata yang tidak memiliki pengaruh atau tidak informatif. *Tokenizing*, yaitu teks pada data-data yang telah diperoleh akan dipisah untuk tiap katanya. *Normalization* adalah mengubah kata-kata gaul dan typo pada kasus ini, menggunakan Bahasa Indonesia. *Stopwords removal* yaitu dengan menghapus kata-kata yang memiliki informasi rendah dari sebuah teks. *Stemming* adalah mencari makna kata dasar. Kemudian setelah data dibersihkan maka disimpan ulang dengan nama databersih dengan ekstensi csv.

Selama tahap ini, data disesuaikan dan dimodifikasi sehingga dapat diproses pada tahap selanjutnya. Tabel 1 menunjukkan contoh hasil dari enam tahap prapemrosesan: *Case Folding*, *Remove Punctuation*, *Tokenizing*, *Normalisasi*, *Filtering*, dan *Stemming*.

<b>Proses</b>	<b>Hasil</b>
<i>Data Collection</i>	Aplikasi gajelas, mau login susahnya mnta ampun
<i>Case Folding</i>	aplikasi gajelas, mau login susahnya mnta ampun
<i>Remove Punctuation</i>	aplikasi gajelas mau login susahnya mnta ampun
<i>Tokenizing</i>	'aplikasi' 'gajelas' 'mau' 'login' 'susahnya' 'mnta' 'ampun'
<i>Normalization</i>	'aplikasi' 'tidak' 'jelas' 'mau' 'login' 'susahnya' 'minta' 'ampun'
<i>Filtering</i>	'aplikasi' 'login' 'susahnya' 'minta' 'ampun'
<i>Stemming</i>	'aplikasi' 'login' 'susah' 'minta' 'ampun'

Tabel 1. Hasil *Preprocessing*

## HASIL DAN PEMBAHASAN

### 1. Analisis Sentimen dengan VADER

Setiap ulasan yang sudah melalui tahapan data preprosesing akan dianalisis menggunakan VADER, yang akan menilai dan memberikan skor sentimen berdasarkan kata-kata dan frasa yang digunakan dalam ulasan tersebut. Skor yang diberikan oleh VADER mencakup tiga nilai utama:

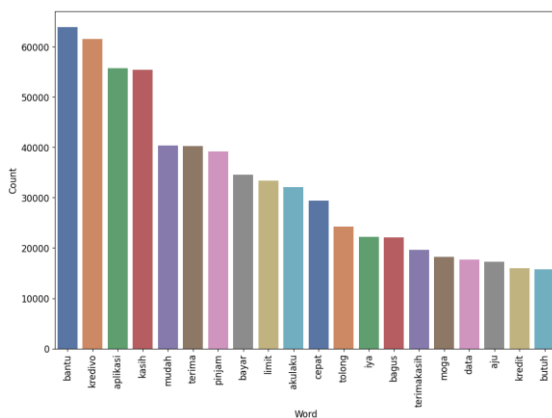
- Positif: menunjukkan jumlah ulasan yang memiliki sentimen positif.
- Negatif: menunjukkan jumlah ulasan yang memiliki sentimen negatif.
- Netral: menunjukkan jumlah ulasan yang memiliki sentimen netral.

Berdasarkan tabel 2, dari analisis yang dilakukan, ditemukan bahwa dari total ulasan yang dikumpulkan, sebanyak 166.938 ulasan dikategorikan sebagai ulasan positif. Ini menunjukkan bahwa mayoritas pengguna memberikan tanggapan yang baik dan memiliki pengalaman positif dengan aplikasi tersebut. Sebaliknya, terdapat 43.202 ulasan negatif, yang menunjukkan adanya sejumlah pengguna yang tidak puas dengan aplikasi atau mengalami masalah tertentu yang mempengaruhi persepsi mereka secara negatif. Selain itu, terdapat 209.551 ulasan netral, yang berarti bahwa sejumlah besar pengguna memberikan ulasan yang tidak menunjukkan sentimen kuat ke arah positif atau negatif. Ulasan netral ini bisa mencerminkan ulasan deskriptif yang lebih informatif tanpa menunjukkan bias sentimen yang jelas. Secara keseluruhan, hasil klasifikasi ini memberikan gambaran yang jelas tentang persepsi pengguna terhadap aplikasi, yang dapat digunakan untuk evaluasi lebih lanjut dan perbaikan kualitas layanan aplikasi tersebut.

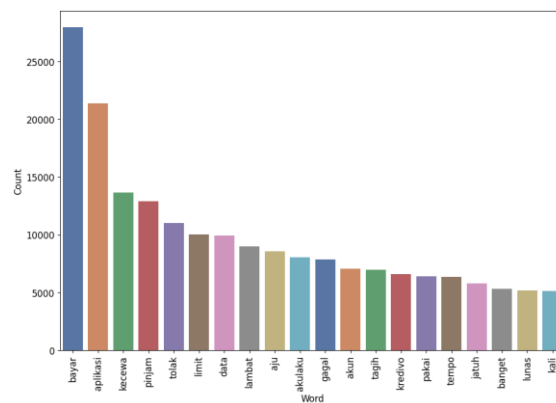
Sentimen	Jumlah Ulasan	Persentase
Positif	166.938	39,78 %
Negatif	43.202	10,29 %
Netral	209.551	49,93 %

Tabel 2. Hasil klasifikasi VADER

Kata yang paling banyak muncul dalam bentuk tabulasi dari data ulasan bersentimen positif dapat dilihat pada gambar 2. Kata yang paling banyak muncul dalam bentuk tabulasi dari data ulasan bersentimen negatif dapat dilihat pada gambar 3.



Gambar 2. Kata paling banyak muncul dari data ulasan bersentimen positif

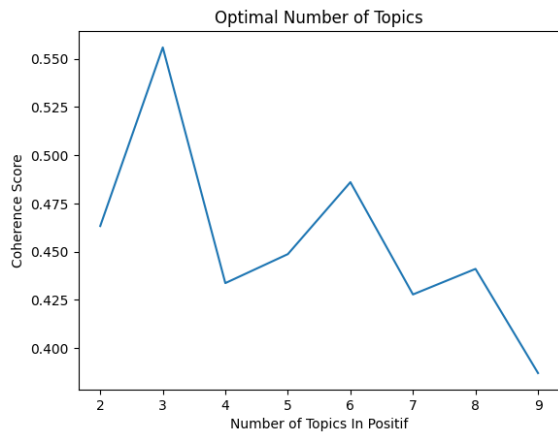


Gambar 3. Kata paling banyak muncul dari data ulasan bersentimen negatif

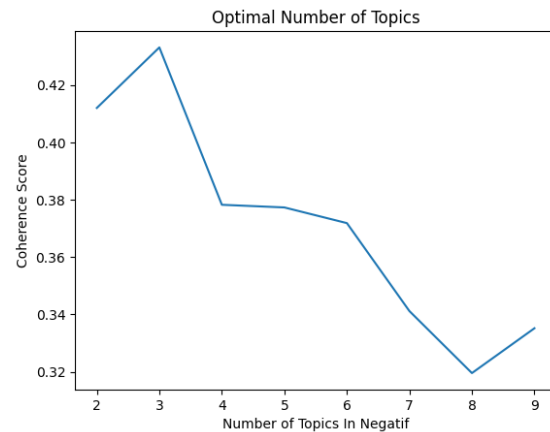
## 2. Topic Modeling dengan LDA

Pada tahapan topic modeling menggunakan Latent Dirichlet Allocation (LDA), penelitian ini memanfaatkan library gensim yang dirancang untuk mengekstraksi topik semantik secara otomatis dari dokumen-dokumen dengan efisien menggunakan software berbasis Python. Penelitian ini terlebih dahulu menentukan kluster words of topic yaitu mengacu pada jumlah kata yang akan digunakan untuk menyusun topik. Jumlah kata ditentukan sebanyak 10 kata, dengan tujuan agar

topik yang dihasilkan tidak bersinggungan satu sama lain. Selanjutnya penelitian ini melakukan evaluasi pemodelan topik dengan menentukan nilai coherence. Parameter yang dijadikan acuan dalam evaluasi ini menggunakan number of topics atau jumlah topik yang diperoleh dalam dokumen. Semakin tinggi nilai coherence, semakin baik juga pemahaman interpretasi manusia terhadap topik tersebut. Untuk mendapatkan topik dengan model terbaik, penelitian ini menganalisis nilai coherence dengan melakukan running sebanyak 50 passes. Hasil dari nilai coherence terdapat pada Gambar 4 untuk pemodelan topik ulasan positif dan Gambar 5 untuk pemodelan topik ulasan negatif.



Gambar 4. Nilai Koherence untuk pemodelan topik sentiment Positif



Gambar 5. Nilai Koherence untuk pemodelan topik sentiment Negatif

Selanjutnya proses pembentukan model LDA dilakukan dengan menggunakan bantuan library gensim. Hasil pemodelan topik menggunakan LDA ditunjukkan pada tabel 2 dan tabel 3. Berdasarkan tabel 2 dan tabel 3, dapat diinterpretasikan 6 topik dari pemodelan dengan data sentiment positif dan negatif. Hasil interpretasi topik digunakan untuk melihat tren ulasan pengguna aplikasi Pinjaman Online.

word	topic_id	importance	word_count	word	topic_id	importance	word_count
bintang	0	0.044908	13.318	aplikasi	0	0.041774	21.363
limit	0	0.037198	33.401	bayar	0	0.020498	27.957
pinjam	0	0.036795	39.175	pinjam	0	0.020025	12.925
kasih	0	0.036793	55.412	tagih	0	0.018930	6.999
bayar	0	0.035203	34.555	tipu	0	0.012616	2.367
iya	0	0.021649	22.201	bunga	0	0.012537	1.959
lambat	0	0.014647	9.733	uang	0	0.010224	2.286
aplikasi	0	0.013914	55.720	banget	0	0.010209	5.338
jt	0	0.013129	5.246	lambat	0	0.009634	8.998
bunga	0	0.013071	15.710	orang	0	0.008920	2.005
data	1	0.051311	17.657	data	1	0.066818	9.951
tolong	1	0.036735	24.272	tolak	1	0.054538	11.026



word	topic_id	importance	word_count
pinjam	1	0.033214	39.175
easy	1	0.029537	5.999
aplikasi	1	0.027918	55.720
hapus	1	0.027639	7.594
cash	1	0.026177	5.971
kasih	1	0.026167	55.412
mohon	1	0.025215	13.711
terima	1	0.023281	40.224
bantu	2	0.085563	63.850
mudah	2	0.058983	40.376
kasih	2	0.057005	55.412
terima	2	0.049344	40.224
easycash	2	0.040788	9.998
cepat	2	0.039172	29.447
moga	2	0.031808	18.233
aplikasi	2	0.029814	55.720
kredivo	2	0.025775	61.570
terimakasih	2	0.022435	19.585

Tabel 2. Kumpulan kata dan id topik hasil pemodelan dengan LDA untuk sentiment Positif

word	topic_id	importance	word_count
aplikasi	1	0.045213	21.363
gagal	1	0.035887	7.856
aju	1	0.031758	8.578
salah	1	0.026569	4.630
hapus	1	0.020517	2.987
pinjam	1	0.018267	12.925
daftar	1	0.016759	3.691
tolong	1	0.013253	3.128
bayar	2	0.084115	27.957
kecewa	2	0.042049	13.670
limit	2	0.035983	10.051
pinjam	2	0.029957	12.925
lambat	2	0.027148	8.998
lunas	2	0.025847	5.211
tempo	2	0.023650	6.374
blokir	2	0.022665	5.027
jatuh	2	0.021272	5.797
aplikasi	2	0.020337	21.363

Tabel 3. Kumpulan kata dan id topik hasil pemodelan dengan LDA untuk sentiment Negatif

Hasil pemodelan topik yang ditampilkan untuk data dengan sentimen positif menunjukkan tiga topik utama yang muncul dari teks yang dianalisis, kemungkinan besar ulasan atau komentar pengguna terkait layanan pinjaman atau aplikasi keuangan:

1. Topik 0 : Pinjaman dan Batasan; Topik ini berfokus pada aspek inti dari layanan pinjaman, seperti: (Bintang: Menunjukkan kemungkinan adanya sistem penilaian atau ulasan pengguna) (Limit: Mengacu pada batasan pinjaman yang diberikan) (Pinjam, Bayar: Menunjukkan proses transaksi pinjaman) (Kasih: Mungkin terkait dengan permintaan atau persetujuan pinjaman) (Lambat: Kemungkinan keluhan tentang proses yang lambat) (Bunga: Mengacu pada biaya pinjaman)
2. Topik 1: Aplikasi dan Data; Topik ini lebih mengarah pada aspek teknis dan pengalaman pengguna aplikasi: (Data: Kemungkinan terkait dengan data pribadi atau informasi yang dibutuhkan dalam aplikasi) (Tolong, Mohon: Menunjukkan permintaan bantuan atau dukungan) (Easy (Cash): Mungkin fitur atau produk yang ditawarkan dalam aplikasi) (Aplikasi, Hapus: Mengacu pada penggunaan dan pengelolaan aplikasi)
3. Topik 3: Kemudahan dan Dukungan; Topik ini mencerminkan sentimen positif terhadap layanan: (Bantu, Mudah: Menunjukkan kemudahan penggunaan dan dukungan yang baik) (Kasih, Terima: Menunjukkan rasa terima kasih atas layanan yang diberikan) (Easycash,

Cepat: Menunjukkan pengalaman positif dengan proses yang cepat dan mudah) (Moga, Terimakasih: Ungkapan harapan dan terima kasih) (Kredivo: Kemungkinan nama aplikasi atau layanan yang dimaksud)

Hasil pemodelan topik yang ditampilkan untuk data dengan sentimen negatif menunjukkan tiga topik utama yang muncul dari teks yang dianalisis:

1. Topik 0: Masalah pada Aplikasi Pinjaman (Topik ini sepertinya berkaitan dengan masalah atau pengalaman negatif seputar aplikasi pinjaman. Kata-kata seperti "bayar", "pinjam", "tagih", "tipu", "bunga", "uang", "lambat", dan "orang" menunjukkan bahwa topik ini mungkin mencakup keluhan tentang proses pembayaran, penagihan, bunga, atau bahkan kemungkinan penipuan)
2. Topik 1: Penolakan Aplikasi dan Masalah Data (Topik ini sepertinya berfokus pada masalah teknis dan penolakan aplikasi. Kata-kata seperti "data", "tolak", "gagal", "aju", "salah", "hapus", "daftar", dan "tolong" mengindikasikan bahwa topik ini mungkin mencakup masalah terkait data pengguna, penolakan aplikasi, atau kesulitan dalam proses pendaftaran)
3. Topik 2: Kekecewaan terhadap Pembayaran dan Limit (Topik ini tampaknya terkait dengan kekecewaan pengguna terhadap aspek pembayaran dan limit pinjaman. Kata-kata seperti "bayar", "kecewa", "limit", "lambat", "lunas", "tempo", "blokir", dan "jatuh tempo" menunjukkan bahwa topik ini mungkin mencakup keluhan tentang keterlambatan pembayaran, limit pinjaman yang rendah, atau masalah terkait blokir akun)

## **KESIMPULAN DAN SARAN**

Penelitian ini menganalisis sentimen dan topik ulasan pengguna pada aplikasi pinjaman online Akulaku, Kredivo, dan EasyCash di Google Play Store menggunakan teknik text mining, analisis sentimen VADER, dan pemodelan topik LDA. Hasil penelitian menunjukkan mayoritas ulasan pengguna bernada positif, mencerminkan kepuasan terhadap aplikasi pinjol. Sentimen positif ini terkait dengan kemudahan penggunaan aplikasi, kecepatan layanan, serta responsivitas dan kualitas dukungan pelanggan. Namun, ada juga ulasan negatif yang menyoroti masalah seperti kesulitan dalam pengajuan pinjaman, masalah teknis pada aplikasi, dan pengalaman buruk terkait penagihan dan pembayaran. Penelitian ini memberikan kontribusi signifikan dalam memahami sentimen dan preferensi pengguna aplikasi pinjol di Indonesia, meskipun memiliki keterbatasan, termasuk cakupan data yang terbatas pada Google Play Store dan potensi keterbatasan dalam analisis sentimen menggunakan VADER. Penelitian lebih lanjut diharapkan dapat memperluas cakupan data dan menggunakan metode analisis sentimen yang lebih canggih untuk mendapatkan pemahaman yang lebih mendalam tentang sentimen pengguna terhadap aplikasi pinjol.

Penelitian ini memiliki beberapa keterbatasan. Pertama, data hanya diambil dari Google Play Store. Penelitian selanjutnya bisa mencakup ulasan dari platform lain seperti App Store atau forum diskusi online untuk mendapatkan gambaran yang lebih lengkap. Penelitian ini menggunakan VADER untuk analisis sentimen, yang mungkin kurang akurat dalam memahami nuansa bahasa Indonesia. Metode analisis yang lebih canggih atau model khusus bahasa Indonesia bisa dipertimbangkan ke depan. Dengan mengatasi keterbatasan ini, penelitian berikutnya bisa memberikan pemahaman yang lebih baik tentang sentimen pengguna terhadap aplikasi pinjol dan memberikan rekomendasi yang lebih tepat bagi penyedia layanan pinjol dan pembuat kebijakan.

## DAFTAR PUSTAKA

- Aditya, Dio Rizki, Endang Supriyati, and Tri Listyorini. 2022. "Analisis Sentimen Pengguna Twitter Terhadap Rokok Elektrik (Vape) Di Indonesia Menggunakan Metode Naïve Bayes." *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)* 7(1): 43–50. doi:10.29100/jipi.v7i1.2145.
- Amir Latif, Rana M., M. Talha Abdullah, Syed Umair Aslam Shah, Muhammad Farhan, Farah Ijaz, and Abdul Karim. 2019. "Data Scraping from Google Play Store and Visualization of Its Content for Analytics." *2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*: 1–8. doi:10.1109/ICOMET.2019.8673523.
- Anggraeni, Windi Ariesti, Fikri Fahru Roji, and Muslim Alkautsar. 2023. "Analisis Sentimen Publik Terhadap Kebijakan Insentif Perpajakan Dengan Pendekatan VADER (Valence Aware Dictionary And Sentiment Reasoner)." *Jurnal Proaksi* 10(4): 465–77. doi:10.32534/jpk.v10i4.4732.
- Barbado, Rodrigo, Oscar Araque, and Carlos A Iglesias. 2019. "A Framework for Fake Review Detection in Online Consumer Electronics Retailers." *Information Processing & Management* 56(4): 1234–44. doi:https://doi.org/10.1016/j.ipm.2019.03.002.
- Carranza, Rocío, Estrella Díaz, Carlos Sánchez-Camacho, and David Martín-Consuegra. 2021. "E-Banking Adoption: An Opportunity for Customer Value Co-Creation." *Frontiers in Psychology* 11. doi:10.3389/fpsyg.2020.621248.
- Chung, Sunghun, Keongtae Kim, Chul Ho Lee, and Wonseok Oh. 2023. "Interdependence between Online Peer-to-peer Lending and Cryptocurrency Markets and Its Effects on Financial Inclusion." *Production and Operations Management* 32(6): 1939–57. doi:10.1111/poms.13950.
- Davis, Bradley M, Samineh C Gillmore, and Derek Millard. 2020. "Sentiment Analysis of Participant Comments in a User Centered Design Study for Degraded Visual Environment Sensor Visualization." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64(1): 2075–78. doi:10.1177/1071181320641502.
- Dieng, Adji B, Joaquín R Vidal, and David M Blei. 2020. "Topic Modeling in Embedding Spaces." *Transactions of the Association for Computational Linguistics*. doi:10.1162/tacl\_a\_00325.
- Dina, Nasa Zata, Riky Tri Yunardi, Aji Akbar Firdaus, and Nyoman Juniarta. 2021. "Measuring User Satisfaction of Educational Service Applications Using Text Mining and Multicriteria Decision-Making Approach." *International Journal of Emerging Technologies in Learning* 16(17): 76–88. doi:10.3991/ijet.v16i17.22939.
- Dong, Manqing, Lina Yao, Xianzhi Wang, Boualem Benatallah, Chaoran Huang, and Xiaodong Ning. 2020. "Opinion Fraud Detection via Neural Autoencoder Decision Forest." *Pattern Recognition Letters* 132: 21–29. doi:https://doi.org/10.1016/j.patrec.2018.07.013.

- Gonzalez, Laura. 2023. "Financial Literacy in For-Profit vs pro-Social Peer-to-Peer Lending." *Managerial Finance* 49(2): 315–37. doi:10.1108/MF-07-2021-0329.
- Gribkov, Egor, and Yuri Yekhlakov. 2020. "Neural Network Model for User Request Analysis during Software Operations and Maintenance Phase." (1 Vol.14): 7–18. <https://bijournal.hse.ru/en/2020--1>.
- Huang, Li, and Henry N. Pontell. 2023. "Crime and Crisis in China's P2P Online Lending Market: A Comparative Analysis of Fraud." *Crime, Law and Social Change* 79(4): 369–93. doi:10.1007/s10611-022-10053-y.
- Jelodar, Hamed, Yongli Wang, Chi Yuan, Xia Feng, Xiahui Jiang, Yanchao Li, and Liang Zhao. 2018. "Latent Dirichlet Allocation (LDA) and Topic Modeling: Models, Applications, a Survey." *Multimedia Tools and Applications*. doi:10.1007/s11042-018-6894-4.
- Jung, Hae Sun, Seon Hong Lee, Haein Lee, and Jang Hyun Kim. 2023. "Predicting Bitcoin Trends Through Machine Learning Using Sentiment Analysis with Technical Indicators." *Computer Systems Science and Engineering* 46(2). doi:10.32604/csse.2023.034466.
- Lee, Seon Hong, Haein Lee, and Jang Hyun Kim. 2022. "Enhancing the Prediction of User Satisfaction with Metaverse Service Through Machine Learning." *Computers, Materials & Continua* 72(3). doi:10.32604/cmc.2022.027943.
- Liu, Siru, and Jialin Liu. 2021. "Public Attitudes toward COVID-19 Vaccines on English-Language Twitter: A Sentiment Analysis." *Vaccine* 39(39): 5499–5505. doi:<https://doi.org/10.1016/j.vaccine.2021.08.058>.
- Long, Suwan (Cheng), Brian Lucey, Ying Xie, and Larisa Yarovaya. 2023. "‘I Just like the Stock’: The Role of Reddit Sentiment in the GameStop Share Rally." *Financial Review* 58(1): 19–37. doi:<https://doi.org/10.1111/fire.12328>.
- Masyhuri, Muhammad. 2022. "Topic Modelling Application for Determining Competitiveness Factors of the Small Business Firms." *International Journal of Social Science and Business*. doi:10.23887/ijssb.v6i2.43164.
- Ming Wang Ge Zhan, Kin Keung Lai Liangbo Zhang, and Lu Meng. 2022. "Posts and Reviews in P2P Online Lending Platforms: A Sentiment Analysis and Cross-Culture Comparison." *Behaviour & Information Technology* 41(16): 3591–97. doi:10.1080/0144929X.2021.2005679.
- Munandar, Adi Ariyo, Farikhin Farikhin, and Catur Edi Widodo. 2023. "Sentimen Analisis Aplikasi Belajar Online Menggunakan Klasifikasi SVM." *JOINTECS (Journal of Information Technology and Computer Science)* 8(2): 77. doi:10.31328/JOINTECS.V8I2.4747.
- Noei, Ehsan, Feng Zhang, and Ying Zou. 2021. "Too Many User-Reviews! What Should App Developers Look at First?" *IEEE Transactions on Software Engineering* 47(2): 367–78. doi:10.1109/TSE.2019.2893171.
- Nofiyani, Nofiyani, and Wulandari Wulandari. 2022. "Implementasi Electronic Data Processing Untuk Meningkatkan Efektifitas Dan Efisiensi Pada Text Mining." *Jurnal Media Informatika Budidarma*. doi:10.30865/mib.v6i3.4332.
- Okuboyejo, Senanu, and Ooreofe Koyejo. 2021. "Examining Users' Concerns While Using Mobile Learning Apps." *International Journal of Interactive Mobile Technologies (iJIM)* 15(15): 47–58. doi:10.3991/ijim.v15i15.22345.
- Puspita, Evi, Diqy Fakhrun Shiddieq, and Fikri Fahru Roji. 2024. "Pemodelan Topik Pada Media Berita Online Menggunakan Latent Dirichlet Allocation (Studi Kasus Merek Somethinc)." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 4(2): 481–89. doi:10.57152/malcom.v4i2.1204.
- Roji, Fikri Fahru, Nava Gia Ginasta, Yayan Cahyan, Dinar Rahayu, and Dendi Ramdani. 2023. "Review Analysis of SatuSehat Application Using Support Vector Machine and Latent

- Dirichlet Allocation Modeling.” *RISTEC : Research in Information Systems and Technology* 4(1): 76–88.
- Ruangkanjanases, Athapol, Pasika Jeebjong, Natalia, and Lim Sanny. 2021. “E-Wom and Its Impacts on Purchasing Behavior: A Comparative Study between Thai and Indonesian Millennials.” *International Journal of Electronic Commerce Studies* 12(1): 65–82. doi:10.7903/IJECS.1893.
- Salsabila, Aqila, Jeremia Jordan Sihombing, and Rizky Irvandi Sitorus. 2022. “Implementasi Algoritma Support Vector Machine Untuk Analisis Sentimen Aplikasi OLX Di Playstore.” *Journal of Informatics and Data Science*. doi:10.24114/j-ids.v1i2.42597.
- Stein, Nicole, Stefan Spinler, Helga Vanthournout, and Vered Blass. 2020. “Consumer Perception of Online Attributes in Circular Economy Activities.” *Sustainability* 12(5). doi:10.3390/su12051914.
- Ullah, Saif, Umar Safdar Kiani, Basharat Raza, and Abdullah Mustafa. 2022. “Consumers’ Intention to Adopt m-Payment/m-Banking: The Role of Their Financial Skills and Digital Literacy.” *Frontiers in Psychology* 13. doi:10.3389/fpsyg.2022.873708.
- Utami, Nengah W, and Made Artana. 2022. “Text Mining Dalam Analisis Sentimen Pembelajaran Daring Di Masa Pandemi Covid 19 Menggunakan Algoritma K-Nearest Neighbor.” *Jurnal Informatika Teknologi Dan Sains*. doi:10.51401/jinteks.v4i2.2034.
- Vijayarathy, Leo R, and Charles W Butler. 2016. “Choice of Software Development Methodologies: Do Organizational, Project, and Team Characteristics Matter?” *IEEE Software* 33(5): 86–94. doi:10.1109/MS.2015.26.
- Wibowo, Ari, Wachid Darmawan, and Nurul Amalia. 2022. “Komparasi Metode Naïve Bayes Dan K-Nearest Neighbor Terhadap Analisis Sentimen Pengguna Aplikasi Pedulilindungi.” *Ic-Tech*. doi:10.47775/icttech.v17i1.234.
- Xia, Yufei, Xinyi Guo, Yinguo Li, Lingyun He, and Xueyuan Chen. 2022. “Deep Learning Meets Decision Trees: An Application of a Heterogeneous Deep Forest Approach in Credit Scoring for Online Consumer Lending.” *Journal of Forecasting* 41(8): 1669–90. doi:https://doi.org/10.1002/for.2891.
- Zhao, Huiying, Huaxin Peng, and Wanqi Li. 2022. “Analysis of Factors Affecting Individuals’ Online Consumer Credit Behavior: Evidence From China.” *Frontiers in Psychology* 13. doi:10.3389/fpsyg.2022.922571.
- Zhuo, Jia. 2022. “Consumer Demand Behavior Mining and Product Recommendation Based on Online Product Review Mining and Fuzzy Sets” ed. Zaoli Yang. *Mathematical Problems in Engineering* 2022: 1216475. doi:10.1155/2022/1216475.
- Zuluaga, Martha. 2023. “Topic Modeling: Perspectives From a Literature Review.” *Ieee Access*. doi:10.1109/access.2022.3232939.