

**ANALISIS SENTIMEN PENGGUNA X TERHADAP LAYANAN *BUY NOW, PAY LATER (BNPL)* DI INDONESIA MENGGUNAKAN
METODE *NAÏVE BAYES CLASSIFIER* DAN *LATENT DIRICHLET ALLOCATION***

Tugas Akhir

Diajukan untuk memenuhi persyaratan mencapai derajat
Sarjana Sistem Informasi



Viola Gabrie Virginia

NPM: 211711393

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2025**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

ANALISIS SENTIMEN PENGGUNA X TERHADAP LAYANAN BUY NOW, PAY LATER (BNPL) DI INDONESIA
MENGUNAKAN METODE NAÏVE BAYES CLASSIFIER DAN LATENT DIRICHLET ALLOCATION

yang disusun oleh

Viola Gabrie Virginia

211711393

dinyatakan telah memenuhi syarat pada tanggal 25 Juli 2025

		Keterangan
Dosen Pembimbing 1	: Julius Galih Prima Negara, S.Kom., S.A.P., M.Kom.	Telah Menyetujui
Dosen Pembimbing 2	: Citra Yuyu' Palangan, S.T., M.Sc.	Telah Menyetujui
Tim Penguji		
Penguji 1	: Julius Galih Prima Negara, S.Kom., S.A.P., M.Kom.	Telah Menyetujui
Penguji 2	: Emanuel Ristian Handoyo, S.T., M.Eng.	Telah Menyetujui
Penguji 3	: Y. Sigit Purnomo WP., S.T., M.Kom., Ph.D.	Telah Menyetujui

Yogyakarta, 25 Juli 2025

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. Ir. Parama Kartika Dewa SP., S.T., M.T.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

LEMBAR PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Viola Gabrie Virginia
NPM : 211711393
Program Studi : Sistem Informasi
Fakultas : Fakultas Teknologi Industri
Judul Penelitian : Analisis Sentimen Pengguna X Terhadap Layanan *Buy Now, Pay Later (BNPL)* di Indonesia Menggunakan Metode *Naïve Bayes Classifier* Dan *Latent Dirichlet Allocation*

Menyatakan dengan ini:

1. Skripsi ini adalah benar merupakan hasil karya sendiri dan tidak merupakan salinan sebagian atau keseluruhan dari karya orang lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta, berupa Hak Bebas Royalti non eksklusif (*Non-Exclusive-Royalty-Free Right*) atas Penelitian ini, dan berhak menyimpan, mengelola dalam pangkalan data, mendistribusikan, serta menampilkan untuk kepentingan akademis, tanpa perlu meminta izin selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum yang mengikuti atas pelanggaran Hak Cipta dalam pembuatan skripsi ini.

Demikianlah pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 7 Juli 2025

Yang menyatakan,



Viola Gabrie Virginia
211711393

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan Yesus Kristus atas kasih karunia, penyertaan, dan kekuatan-Nya yang tiada henti yang memungkinkan penulis menyelesaikan skripsi berjudul “Analisis Sentimen Pengguna X Terhadap Layanan *Buy Now, Pay Later (BNPL)* di Indonesia Menggunakan Metode *Naïve Bayes Classifier* Dan *Latent Dirichlet Allocation*” ini dengan baik. Tanpa pertolongan dan penyertaan-Nya, perjalanan ini tidak mungkin dapat ditempuh dengan segala keterbatasan dan tantangan yang ada. Skripsi ini disusun sebagai salah satu syarat akademis untuk memperoleh gelar Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa penyusunan skripsi ini bukanlah hal yang mudah. Tantangan datang dari berbagai sisi, baik teknis, emosional, maupun waktu. Namun dengan dukungan, doa, dan semangat dari berbagai pihak, penulis dapat melalui setiap proses hingga mencapai tahap ini. Oleh karena itu, dengan penuh rasa syukur dan hormat, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Papa dan Mama tercinta, atas doa yang tak pernah putus, kasih sayang, dukungan moral maupun materi, serta semangat yang selalu diberikan di setiap langkah perjalanan studi ini. Terima kasih sudah menjadi rumah yang paling nyaman dan tempat kembali yang penuh kehangatan.
2. Bapak Julius Galih Prima Negara, S.Kom., S.A.P., M.Kom dan Ibu Citra Yuyu' Palangan, S.T., M.Sc. selaku dosen pembimbing skripsi penulis, atas waktu, perhatian, arahan, serta kesabaran yang luar biasa dalam membimbing penulis selama proses penyusunan skripsi ini.
3. Seluruh dosen Sistem Informasi Universitas Atma Jaya Yogyakarta, yang telah membagikan ilmu, mendukung, serta membimbing penulis selama masa perkuliahan hingga penulis bisa berada di titik ini.
4. Sahabat, teman-teman, dan *partner* penulis yang senantiasa memberikan dukungan, doa, serta meluangkan waktu di tengah kesibukan masing-masing. Terima kasih telah hadir menjadi penguat di setiap langkah proses

ini. Kehadiran kalian menjadi bagian tak terpisahkan dari perjalanan yang penuh makna ini.

5. Terakhir, penulis mengucapkan terima kasih kepada diri sendiri. Terima kasih telah memilih untuk terus melangkah, bahkan saat rasanya ingin berhenti. Terima kasih telah bertahan sejauh ini, tetap percaya di tengah keraguan, dan tidak menyerah di masa-masa tersulit. Perjalanan ini mungkin tidak selalu mudah, tapi langkah demi langkah telah membawa sampai di titik ini — dan itu sudah sangat cukup. Terima kasih, sudah tidak menyerah.

“Segala perkara dapat kutanggung di dalam Dia yang memberi kekuatan kepadaku”

Filipi 4:13

INTISARI

Layanan *Buy Now, Pay Later (BNPL)* menjadi salah satu bentuk inovasi dalam sistem pembayaran digital yang semakin banyak dibicarakan di media sosial. Namun, keberagaman opini dan topik yang dibagikan oleh pengguna X belum sepenuhnya dipahami oleh regulator, sehingga berpotensi menimbulkan kesenjangan dalam penyusunan kebijakan perlindungan konsumen dan pengawasan terhadap layanan ini. Penelitian ini bertujuan untuk mengklasifikasikan opini pengguna terhadap layanan BNPL ke dalam tiga kelas sentimen—positif, netral, dan negatif—serta mengidentifikasi topik-topik utama yang dibahas menggunakan algoritma Naïve Bayes dan Latent Dirichlet Allocation (LDA). Proses penelitian meliputi pengumpulan data, *preprocessing data*, klasifikasi sentimen, dan pemodelan topik. Hasil pengujian klasifikasi sentimen menunjukkan bahwa metode *oversampling* memberikan kinerja terbaik dengan akurasi sebesar 85,42%, *precision* 88%, *recall* 85%, dan *F1-score* sebesar 86%. Sementara itu, pemodelan topik menghasilkan 8 topik pada masing-masing kelas sentimen, dengan nilai koherensi tertinggi untuk sentimen positif sebesar 0,517, sentimen netral 0,543, dan negatif sebesar 0,419. Temuan ini menunjukkan bahwa kombinasi metode Naïve Bayes dan LDA efektif dalam menggali persepsi dan isu terkait layanan BNPL di media sosial, yang dapat menjadi dasar pertimbangan regulator dalam merancang kebijakan yang lebih responsif dan adaptif terhadap kebutuhan masyarakat.

Kata Kunci: BNPL; Analisis Sentimen; *Naïve Bayes Classifier*; Pemodelan Topik; *Latent Dirichlet Allocation*

ABSTRACT

Buy Now, Pay Later (BNPL) services have emerged as a form of innovation in digital payment systems and are increasingly discussed on social media. However, the diversity of opinions and topics shared by users on platform X has not been fully understood by regulators, potentially leading to a gap in consumer protection policies and oversight mechanisms. This study aims to classify user opinions on BNPL services into three sentiment categories—positive, neutral, and negative—and to identify the main discussion topics using the Naïve Bayes algorithm and Latent Dirichlet Allocation (LDA). The research process includes data collection, data preprocessing, sentiment classification, and topic modeling. Sentiment classification results show that the oversampling method yields the best performance, with an accuracy of 85.42%, precision of 88%, recall of 85%, and an F1-score of 86%. Meanwhile, the topic modeling process produces 8 dominant topics in each sentiment category, with the highest coherence scores of 0.517 for positive sentiment, 0.543 for neutral, and 0.419 for negative. These findings demonstrate that the combination of Naïve Bayes and LDA is effective in extracting public perceptions and emerging issues related to BNPL services on social media, providing valuable insights for regulators in designing more responsive and adaptive policies to meet public needs.

Keywords: BNPL; Sentiment Analysis; Naïve Bayes Classifier; Topic Modeling; Latent Dirichlet Allocation

DAFTAR ISI

HALAMAN PENGESAHAN	ii
LEMBAR PERNYATAAN KEASLIAN PENELITIAN	iii
PRAKATA	iv
INTISARI.....	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Pertanyaan Penelitian.....	5
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
1.6. Batasan Masalah.....	6
1.7. Bagan Keterkaitan.....	7
BAB II TINJAUAN PUSTAKA	8
2.1. Studi Sebelumnya	8
2.2. Dasar Teori.....	13
2.2.1 <i>Buy Now, Pay Later (BNPL)</i>	13
2.2.2 <i>Natural Language Processing (NLP)</i>	13
2.2.3 <i>Text Mining</i>	13
2.2.4 Analisis Sentimen	14
2.2.5 <i>Preprocessing Text</i>	14
2.2.6 Klasifikasi Sentimen.....	15
2.2.7 <i>Topic Modeling</i>	17
BAB III METODOLOGI PENELITIAN	18
3.1. Tahapan Penelitian	18
3.1.1 Identifikasi Masalah	19
3.1.2 Studi Literatur	19
3.1.3 Pengumpulan Data.....	19
3.1.4 <i>Preprocessing Data</i>	20
3.1.5 Pelabelan Sentimen	21
3.1.6 Ekstraksi Fitur	22
3.1.7 Klasifikasi Sentimen.....	23
3.1.8 Pemodelan Topik.....	27
3.1.9 Interpretasi Hasil.....	32

3.1.10	Penarikan Kesimpulan.....	32
3.2.	Alat yang Digunakan	33
BAB IV HASIL DAN PEMBAHASAN		34
4.1.	Pengumpulan Data	34
4.2.	<i>Preprocessing Data</i>	37
4.2.1.	<i>Cleaning</i>	37
4.2.2.	<i>Case Folding</i>	42
4.2.3.	<i>Normalization</i>	43
4.2.4.	<i>Tokenizing</i>	44
4.2.5.	<i>Stopword Removal</i>	45
4.2.6.	<i>Stemming</i>	46
4.3.	Pelabelan Sentimen	47
4.4.	Ekstraksi Fitur	49
4.5.	Klasifikasi Sentimen	51
4.5.1.	<i>Splitting Data</i>	51
4.5.2.	Pelatihan Model	52
4.5.3.	Pengujian Model	53
4.5.4.	Evaluasi Model	57
4.5.5.	Interpretasi Hasil Klasifikasi Sentimen	60
4.6.	Pemodelan Topik	70
4.6.1.	Pemisahan Data Berdasarkan Label Sentimen.....	70
4.6.2.	Penyusunan <i>Corpus</i> dan <i>Dictionary</i>	71
4.6.3.	Penentuan Jumlah Topik Optimal	72
4.6.4.	Pelatihan Model	73
4.6.5.	Penamaan Topik.....	78
4.6.6.	Interpretasi Hasil Pemodelan Topik.....	80
4.7.	Pembahasan Hasil Penelitian	86
4.7.1.	Keterkaitan Hasil Klasifikasi Sentimen dan Pemodelan Topik	86
4.7.2.	Implikasi dan Temuan Penelitian	89
BAB V KESIMPULAN DAN SARAN		91
5.1.	Kesimpulan	91
5.2.	Saran	91
DAFTAR PUSTAKA		93
LAMPIRAN.....		102

DAFTAR GAMBAR

Gambar 1. 1 Bagan Keterkaitan	7
Gambar 3. 1. Tahapan Penelitian	18
Gambar 3. 2. Alur <i>Preprocessing Data</i>	20
Gambar 3. 3. Alur Klasifikasi Sentimen	24
Gambar 3. 4. Konsep Dasar <i>Topic Modeling</i>	28
Gambar 3. 5. <i>Graphical Model</i> LDA	29
Gambar 3. 6. Alur Pemodelan Topik.....	30
Gambar 4. 1. Nilai Sentimen <i>Lexicon</i>	47
Gambar 4. 2. Top 20 Kata dengan Nilai TF-IDF Tertinggi.....	50
Gambar 4. 3. Evaluasi Kinerja Model.....	58
Gambar 4. 4 <i>Confusion Matrix – Oversampling</i>	59
Gambar 4. 5 Distribusi Prediksi Sentimen	59
Gambar 4. 6 WordCloud Sentimen Positif.....	62
Gambar 4. 7. <i>WordCloud</i> Sentimen Netral	64
Gambar 4. 8. <i>WordCloud</i> Sentimen Negatif	67
Gambar 4. 9. Nilai Koherensi Sentimen Positif.....	72
Gambar 4. 10. Nilai Koherensi Sentimen Netral.....	72
Gambar 4. 11. Nilai Koherensi Sentimen Negatif	73

DAFTAR TABEL

Tabel 2. 1. Tabel Perbandingan Penelitian	11
Tabel 4. 1. Hasil Pengumpulan Data	34
Tabel 4. 2. Hasil <i>Cleaning</i>	41
Tabel 4. 3. Hasil <i>Case Folding</i>	43
Tabel 4. 4. Hasil <i>Normalization</i>	44
Tabel 4. 5. Hasil <i>Tokenizing</i>	45
Tabel 4. 6. Hasil <i>Stopword Removal</i>	46
Tabel 4. 7. Hasil <i>Stemming</i>	47
Tabel 4. 8. Jumlah Data Setiap Kelas Sentimen	48
Tabel 4. 9. Tabel Perbandingan <i>Baseline</i> dengan <i>Resampling</i>	54
Tabel 4. 10. <i>Tweet</i> Sentimen Positif	62
Tabel 4. 11. <i>Tweet</i> Sentimen Netral	65
Tabel 4. 12. <i>Tweet</i> Sentimen Negatif	68
Tabel 4. 13. Jumlah Topik Optimal Tiap Sentimen	73
Tabel 4. 14. Tabel Daftar Topik Sentimen Positif	74
Tabel 4. 15. Tabel Daftar Topik Sentimen Netral	75
Tabel 4. 16. Tabel Daftar Topik Sentimen Negatif	77
Tabel 4. 17. Tabel Nama Topik Sentimen Positif	78
Tabel 4. 18. Tabel Nama Topik Sentimen Netral	79
Tabel 4. 19. Tabel Nama Topik Sentimen Negatif	79