ANALISIS SENTIMEN KEPUASAN PELANGGAN USAHA MIKRO, KECIL, DAN MENENGAH MENGGUNAKAN ALGORITMA NAIVE BAYES PADA TEKS BERBAHASA DAERAH MANADO: STUDI KASUS CV TALONGKA-JAYA

Gerrard William Senduk[1], Audyani A. O. Maker[2], Jizreel H. Dumas[3], Marike A. S. Kondoj[4]
[1-4]Program Studi Teknik Informatika, Jurusan Teknik Elektro, Politeknik Negeri Manado, Sulawesi Utara, Indonesia
[1] wiliamsenduk7@gmail.com, [2]audyani123@gmail.com, [3]jizreelhiskiadumas@gmail.com, [4]marikekondoj@polimdo.ac.id

ABSTRACT

Customer reviews are one of the most valuable sources of feedback for small businesses, but most sentiment analysis systems were not designed to handle regional dialect text. In North Sulawesi, Indonesia, customers of local furniture businesses typically write reviews in Manado Malay dialect (Bahasa Melayu Manado/BMM), using words like "nda" (not), "pe" (very), "so" (already), and "gaga" (good/attractive) that standard Indonesian NLP libraries simply cannot recognize. This research builds and evaluates a web-based sentiment analysis system for CV Talongka jaya, a furniture MSME in Manado, using a Multinomial Naive Bayes Classifier combined with a custom BMM-to-Indonesian lexicon normalization module as the primary contribution. The system was developed using Python with Scikit-learn and CountVectorizer, and Laravel for the web dashboard. A dataset of 250 manually labeled customer reviews was evaluated using both an 80/20 train-test split and 10-Fold Cross Validation. System A (with normalization) achieved 84.00% accuracy, 88.89% precision, 82.76% recall, and 85.71% F1-Score, compared to System B (without normalization) which reached only 66.00% accuracy - an 18-percentage-point improvement. The ablation study confirmed that normalization contributed the single largest performance gain among all preprocessing stages.

Keywords: Sentiment Analysis; Naive Bayes; Manado Malay Dialect; Text Normalization; Lexicon; MSME; CRISP-DM

ABSTRAK

Ulasan pelanggan adalah salah satu sumber umpan balik paling berharga bagi usaha kecil, namun sebagian besar sistem analisis sentimen tidak dirancang untuk menangani teks berbahasa daerah. Di Sulawesi Utara, Indonesia, pelanggan usaha furnitur lokal umumnya menulis ulasan dalam dialek Bahasa Melayu Manado (BMM), menggunakan kata-kata seperti "nda" (tidak), "pe" (sangat), "so" (sudah), dan "gaga" (bagus/keren) yang tidak dapat dikenali oleh pustaka NLP Bahasa Indonesia standar. Penelitian ini membangun dan mengevaluasi sistem analisis sentimen berbasis web untuk CV Talongka jaya, sebuah UMKM furnitur di Manado,

menggunakan Multinomial Naive Bayes Classifier yang dikombinasikan dengan modul normalisasi leksikon BMM-ke-Indonesia sebagai kontribusi utama. Sistem dikembangkan menggunakan Python dengan Scikit-learn dan CountVectorizer untuk pipeline machine learning, serta Laravel untuk dashboard web. Dataset berisi 250 ulasan pelanggan berlabel yang dievaluasi menggunakan pembagian data 80/20 dan 10-Fold Cross Validation. Sistem A (dengan normalisasi) mencapai akurasi 84,00%, presisi 88,89%, recall 82,76%, dan F1-Score 85,71% dibandingkan Sistem B (tanpa normalisasi) yang hanya mencapai akurasi 66,00% - peningkatan 18 poin persentase. Ablation study mengonfirmasi bahwa tahap normalisasi memberikan peningkatan performa terbesar di antara semua tahap preprocessing.

Kata Kunci: Analisis Sentimen; Naive Bayes; Bahasa Melayu Manado; Normalisasi Teks; Leksikon; UMKM; CRISP-DM

A. Pendahuluan

Menjalankan usaha furnitur di Manado saat ini berarti menghadapi aliran ulasan pelanggan yang masuk melalui Google Maps, WhatsApp, dan marketplace online. Tantangannya adalah sebagian besar ulasan tersebut tidak ditulis dalam Bahasa Indonesia baku - melainkan dalam dialek Bahasa Melayu Manado (BMM), bahasa sehari-hari masyarakat Sulawesi Utara. Kata-kata seperti "nda" (tidak), "pe" (sangat), dan "gaga" (keren/berkualitas) sangat umum dijumpai dalam ulasan tersebut, namun pada dasarnya tidak terdeteksi oleh sistem NLP mana pun yang dilatih pada teks Bahasa Indonesia standar.

Hal ini merupakan masalah nyata bagi CV Talongka jaya, studi kasus dalam penelitian ini. Tanpa cara otomatis untuk memahami apa yang disampaikan pelanggan, pemilik usaha harus membaca setiap ulasan secara manual - yang semakin tidak realistis seiring pertumbuhan bisnis. Tujuan penelitian ini adalah membangun sistem analisis sentimen yang benar-benar berfungsi untuk teks BMM, bukan hanya untuk Bahasa Indonesia formal.

Penelitian sebelumnya tentang analisis sentimen dalam konteks Manado secara konsisten mengabaikan masalah dialek ini. Petronela (2023) menerapkan Naive Bayes pada ulasan hotel di Kota Manado dan mencapai akurasi 76,20%; Wongkar (2020) menggunakan algoritma yang sama pada data Twitter Manado dan memperoleh 75,58%; Haris (2025) mencapai 65,00% pada ulasan aplikasi UMKM. Tidak satu pun dari penelitian ini yang mencoba

menangani dialek BMM secara eksplisit - semuanya menggunakan pipeline preprocessing Bahasa Indonesia standar pada teks yang setidaknya sebagian ditulis dalam bahasa yang tidak pernah dirancang untuk diproses oleh pipeline tersebut.

Kontribusi inti penelitian ini adalah modul normalisasi leksikon dialek Bahasa Melayu Manado - tabel pencarian berisi 120 token BMM yang dipetakan ke padanan Bahasa Indonesia standarnya, diterapkan sebagai langkah preprocessing sebelum classifier Naive Bayes memproses teks. Sepengetahuan penulis, ini adalah penelitian pertama yang secara eksplisit mengintegrasikan normalisasi dialek BMM ke dalam pipeline analisis sentimen untuk ulasan pelanggan UMKM di Indonesia.

B. Tinjauan Pustaka

1. Analisis Sentimen untuk Konteks UMKM

Analisis sentimen adalah proses komputasi untuk mengidentifikasi dan mengkategorikan opini yang diekspresikan dalam teks, dengan tujuan menentukan sikap penulis terhadap suatu subjek sebagai positif, negatif, atau netral (Liu, 2020). Bagi usaha kecil, analisis sentimen menyediakan cara untuk memproses umpan balik pelanggan dalam skala besar yang tidak mungkin dilakukan secara manual. Jumawan dkk. (2024) menemukan bahwa ulasan online memengaruhi hingga 75,4% keputusan pembelian di marketplace Indonesia, menjadikan analisis sentimen sebagai alat intelijen bisnis yang langsung dapat ditindalanjuti.

2. Dialek Bahasa Melayu Manado (BMM) dalam NLP

BMM bukan sekadar Bahasa Indonesia informal - ia memiliki fitur gramatikal dan leksikal yang berbeda yang dibentuk oleh berabad-abad kontak dengan bahasa Belanda, Portugis, dan bahasa Minahasa lokal (Sneddon, 1983). Untuk keperluan NLP, fitur yang paling berdampak adalah: negasi menggunakan "nda", "nda" alih-alih "tidak" standar; intensifikasi derajat menggunakan "pe" (sangat) dan "kali" (sekali); penandaan aspek menggunakan "so" (sudah); dan kata sifat bermuatan sentimen unik seperti "gaga" (bagus), "alus" (halus), dan kata emosional seperti "pastiu" (kesal). Karena tidak satu pun dari kata-kata ini muncul dalam sumber daya NLP Bahasa Indonesia standar, kata-kata tersebut

pada dasarnya tidak terdeteksi oleh pipeline klasifikasi konvensional.

3. Normalisasi Teks untuk Dialek Rendah Sumber Daya

Normalisasi teks mengonversi teks non-standar ke bentuk terstandarisasi yang dapat diproses oleh alat NLP hilir (Naseem dkk., 2024). Untuk dialek tanpa korpus beranotasi besar, pemetaan token berbasis leksikon - mempertahankan kamus pasangan kata non-standar-ke-standar - adalah pendekatan yang paling praktis. Pendekatan ini transparan, dapat diinterpretasikan, dan tidak memerlukan data pelatihan tersendiri.

4. Naive Bayes untuk Klasifikasi Teks

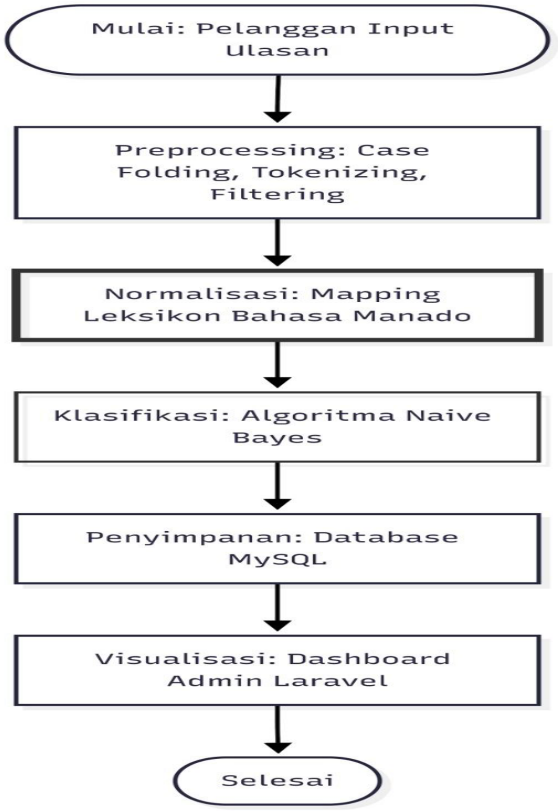
Multinomial Naive Bayes Classifier menghitung probabilitas posterior suatu kelas berdasarkan fitur kata dokumen, dengan asumsi penyederhanaan independensi fitur bersyarat. Terlepas dari asumsi ini, NBC secara konsisten memberikan hasil kompetitif dalam klasifikasi teks, terutama dalam skenario dataset kecil (McCallum & Nigam, 1998). Tabel 1 merangkum studi terkait yang paling relevan dengan penelitian ini.

Tabel 1. Studi Terkait Analisis Sentimen dalam Konteks Manado dan Indonesia

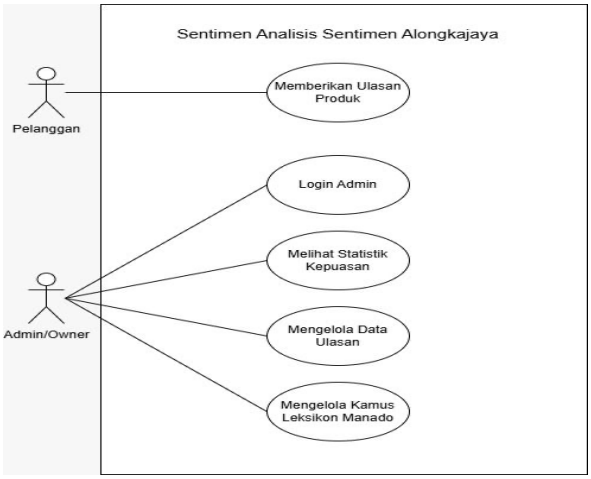
Penulis	Tahun	Konteks	Metode	Akurasi	Norm. BMM?
Petronela	2023	Ulasan hotel, Manado	Naive Bayes	76,20%	Tidak
Wongkar	2020	Twitter, Manado	Naive Bayes	75,58%	Tidak
Haris	2025	Ulasan Aplikasi UMKM	NBC + TF-IDF	65,00%	Tidak
Naseem dkk.	2024	NLP rendah sumber daya	ML + Normalisasi	+20% gain	Ya (umum)
Nguyen dkk.	2025	Dialek Vietnam	Seq2Seq	F1: 0,63	Ya (neural)
Penelitian Ini	2026	UMKM Furnitur,	NBC + Leksikon BMM	84,00%	Ya (120 entri)

Man

ado



Gambar 1. Flowchart Sistem



Gambar 2. Usecase Sistem

B. Metode Penelitian

Penelitian ini mengikuti kerangka CRISP-DM (Cross-Industry Standard Process for Data Mining), model proses yang paling banyak diadopsi untuk proyek machine learning dan penambangan data terapan (Wirth & Hipp, 2000). CRISP-DM terdiri dari enam fase: Pemahaman Bisnis, Pemahaman Data, Persiapan Data, Pemodelan, Evaluasi, dan Penerapan.

1. Pemahaman Bisnis (Fase 1)

Permasalahan bisnis sudah jelas: CV Talongka jaya menerima ulasan pelanggan melalui saluran digital tetapi tidak memiliki cara untuk menganalisisnya secara sistematis. Tujuan sistem adalah mengklasifikasikan setiap ulasan yang masuk sebagai positif atau negatif secara otomatis, lalu menyajikan wawasan melalui dashboard yang dapat digunakan pemilik tanpa pengetahuan teknis. Tantangan utama yang diidentifikasi dalam fase ini adalah kesenjangan linguistik antara apa yang ditulis pelanggan BMM dan apa yang dapat dibaca oleh alat NLP standar.

2. Pemahaman dan Pengumpulan Data (Fase 2)

Sebanyak 250 teks ulasan pelanggan dikumpulkan dari titik kontak digital CV Talongka jaya: Google Maps, thread umpan balik

grup WhatsApp, dan formulir web yang dapat diakses melalui kode QR di toko fisik. Ulasan dikumpulkan selama periode observasi tiga bulan dan merepresentasikan bahasa pelanggan yang alami dan tidak disunting - termasuk kosakata BMM, ejaan informal, dan alih kode antara BMM dan Bahasa Indonesia standar.

Dua anotator, keduanya penutur asli BMM, memberi label setiap ulasan menggunakan protokol yang jelas: Positif jika ulasan mengungkapkan kepuasan terhadap kualitas produk, keahlian, waktu pengiriman, atau layanan; Negatif jika mengandung keluhan tentang cacat, keterlambatan, layanan buruk, atau ketidakpuasan harga. Ketidaksepakatan diselesaikan melalui diskusi. Kesepakatan antar-anotator diukur menggunakan Cohen's Kappa ($\kappa = 0{,}82$), menunjukkan kesepakatan yang kuat (Landis & Koch, 1977).

Tabel 2. Distribusi Dataset Setelah Pembagian Stratifikasi 80/20

Kelas Sentimen	Set Pelatihan (80%)	Set Uji (20%)	Total
Positif	116 ulasan (58%)	29 ulasan (58%)	145 ulasan (58%)
Negatif	84 ulasan (42%)	21 ulasan (42%)	105 ulasan (42%)
Total	200 ulasan	50 ulasan	250 ulasan

3. Persiapan Data - Pipeline Preprocessing (Fase 3)

Semua ulasan diproses melalui pipeline preprocessing empat tahap sebelum klasifikasi. Keempat tahap tersebut adalah:

(1) Case Folding - mengonversi semua teks ke huruf kecil sehingga "Gaga" dan "gaga" diperlakukan sebagai token yang sama. (2) Tokenisasi - memisahkan teks ulasan menjadi token kata individual. (3) Normalisasi Leksikon BMM (Kontribusi Baru) - memetakan setiap token dialek BMM ke padanan Bahasa Indonesia standarnya menggunakan kamus leksikon Manado. Langkah ini dijalankan sebelum filter stopword untuk memastikan kata-kata yang dipetakan ke standar dapat dievaluasi dengan benar. (4) Filter Stopword - menghapus kata fungsi tanpa nilai sentimen menggunakan daftar stopword Bahasa Indonesia kustom.

4. Kamus Leksikon Manado

Leksikon dibangun menggunakan tiga metode: (1) mengekstrak token BMM dari korpus ulasan yang dikumpulkan yang tidak

dapat diproses oleh alat NLP Bahasa Indonesia standar, (2) berkonsultasi dengan lima penutur asli BMM untuk memverifikasi akurasi semantik setiap pemetaan, dan (3) melakukan referensi silang dengan literatur akademis tentang linguistik Bahasa Melayu Manado (Sneddon, 1983). Leksikon yang dihasilkan berisi 120 pasang kata dialek-ke-standar unik yang diorganisasikan dalam enam kategori linguistik.

Tabel 3. Contoh Entri dari Leksikon Bahasa Melayu Manado (total 120 entri)

No.	Token BMM	B. Indonesia Standar	Arti Bahasa Inggris	Kategori
1	nda	tidak	no / not	Partikel negasi
2	blum	tidak / belum	not / not yet	Partikel negasi
3	pe	sangat	very (intensifier)	Intensifier derajat
4	kali	sekali	extremely	Intensifier derajat
5	so	sudah	already	Penanda aspek
6	gaga	bagus / keren	good / attractive	Kata sifat kualitas
7	alus	halus / licin	smooth / polished	Kata sifat kualitas
8	pastiu	kesal / kecewa	annoyed / disappointed	Ekspresi emosional
9	torang	kami / kita	we / us	Kata ganti
10	mar	tetapi	but	Konjungsi
11	kwa	saja	just / only	Partikel
12	mantap skali	sangat bagus	excellent	Kata sifat majemuk

5. Model Klasifikasi (Fase 4)

Classifier yang digunakan adalah Multinomial Naive Bayes (MultinomialNB), diimplementasikan melalui Scikit-learn. CountVectorizer dengan bigram (ngram_range=(1,2)) digunakan untuk ekstraksi fitur - ini adalah pilihan desain yang disengaja, karena MultinomialNB dirancang untuk fitur jumlah kata diskrit. Laplace smoothing (alpha=1.0) menangani token frekuensi nol untuk kata yang muncul di set uji tetapi tidak di data pelatihan. Pipeline terlatih diserialisasi ke file pickle dan dilayani melalui

endpoint API Flask yang dipanggil oleh backend Laravel.

Dasar matematis classifier mengikuti Teorema Bayes: P(C|D) = [P(C) × P(D|C)] / P(D). Probabilitas kata diperhalus menggunakan Laplace: P(w|C) = [count(w,C) + α] / [count(all_w,C) + α × |V|]

6. Metodologi Evaluasi (Fase 5)

Dua strategi evaluasi digunakan. Pertama, pembagian train-test stratifikasi 80/20 mengevaluasi performa pada 50 ulasan yang ditahan. Kedua, 10-Fold Cross Validation pada seluruh dataset 250 ulasan memberikan estimasi akurasi yang lebih stabil dan tidak bergantung pada pembagian. Metrik performa diturunkan dari Matriks Konfusi: Akurasi = (TP+TN)/(TP+TN+FP+FN); Presisi = TP/(TP+FP); Recall = TP/(TP+FN); F1-Score = 2×(Presisi×Recall)/(Presisi+Recall). Ablation study dengan empat konfigurasi pipeline juga dilakukan untuk mengisolasi kontribusi setiap tahap preprocessing.

C. Hasil Penelitian dan Pembahasan

1. Analisis Leksikon

Leksikon Manado dengan 120 entri terdistribusi dalam enam kategori linguistik, seperti ditunjukkan pada Tabel 4. Partikel negasi membentuk bagian terbesar (28%) karena merupakan token paling kritis untuk polaritas sentimen - sistem yang tidak mengenali "nda" sebagai "tidak" akan sepenuhnya melewatkan sinyal negasi. Intensifier derajat mengikuti (24%), yang mempengaruhi kekuatan sentimen. Kata sifat kualitas dan ekspresi emosional (33% gabungan) membawa nilai sentimen langsung dan sangat penting untuk klasifikasi ulasan furnitur.

Tabel 4. Distribusi Kategori Leksikon Manado dan Dampak Sentimen

No.	Kategori	Jumlah	Proporsi	Dampak Sentimen
1	Partikel Negasi (nda, tra, nda)	34	28%	Kritis - membalikkan polaritas kalimat jika terlewat
2	Intensifier Derajat (pe, kali, talalu)	29	24%	Tinggi - memperkuat kekuatan sinyal sentimen
3	Kata Sifat Kualitas	22	18%	Tinggi - sinyal langsung

	(gaga, alus, manta p)			kualitas produk
4	Ekspresi Emosional (pastiu, bagaroh)	18	15%	Sedang - mencerminkan sentimen pengalaman pelanggan
5	Kata Ganti & Referensi (torang, ngoni)	10	8%	Rendah - peran struktural
6	Konjungsi & Partikel (mar, kwa)	7	7%	Rendah - penghubung sintaktis
Total	6 kategori	120	100%	-

2. Transformasi Pipeline Preprocessing

Tabel 5 menunjukkan bagaimana lima ulasan pelanggan aktual dari korpus CV Talongka jaya ditransformasi melalui setiap tahap pipeline preprocessing. Ini membuat efek modul normalisasi menjadi konkret dan dapat diamati.

Tabel 5. Transformasi Pipeline Preprocessing - Sampel Ulasan dari CV Talongka jaya

No.	Ulasan Asli (BMM)	Setelah Case Fold & Tokenisasi	Setelah Normalisasi BMM	Label Benar	Hasil
1	Barang pegaga, nda ada cacat kwa!	barang pegaga nda ada cacat kwa	barang sangat bagus tidak ada cacat saja	Positif	✓ Benar
2	So terlambat dikirim, pastiu torang tunggu	so terlambat dikirim pastiu torang tunggu	sudah terlambat dikirim kesal kami tunggu	Negatif	✓ Benar
3	Kualitas nda bagus, mar harga kwa murah	kualitas nda bagus mar harga kwa murah	kualitas tidak bagus tetapi harga saja murah	Negatif	✓ Benar

4	Furniture pealus, mantap skali buatannya	furniture pealus mantap skali buatannya	furniture sangat halus sangat bagus buatannya	Positif	✓ Benar
5	Pelayanan nda baik, torang nda mau balik	pelayanan nda baik torang nda mau balik	pelayanan tidak baik kami tidak mau kembali	Negatif	✓ Benar

3. Ablation Study

Ablation study menguji setiap tahap preprocessing secara independen dengan melatih dan mengevaluasi empat konfigurasi pipeline yang berbeda pada set uji yang sama berisi 50 ulasan. Ini mengisolasi kontribusi setiap tahap dan mengonfirmasi bahwa peningkatan performa tidak bersifat kebetulan.

Tabel 6. Ablation Study: Performa di Setiap Tahap Preprocessing (Set Uji = 50 ulasan)

Konfigurasi	Deskripsi Pipeline	Akurasi	Presisi	Recall	F1-Score
A	Hanya NBC - tanpa preprocessing	62,00%	63,16 %	63,16 %	63,16 %
B	A + Case Folding & Tokenisasi	66,00%	66,67 %	68,97 %	67,80 %
C	B + Penghapusan Stopword	74,00%	76,47 %	75,86 %	76,16 %
D	C + Normalisasi Leksikon BMM (Sistem Diusulkan)	84,00%	88,89 %	82,76 %	85,71 %

Ablation study mengungkapkan pola yang jelas dan konsisten: setiap tahap preprocessing berkontribusi secara bermakna terhadap performa, dan langkah normalisasi (Konfigurasi C ke D) memberikan lompatan terbesar -

peningkatan akurasi sebesar +10,00 poin persentase. Kontribusi normalisasi yang lebih besar mengonfirmasi bahwa token dialek BMM merupakan sumber utama kesalahan klasifikasi dalam baseline.

4. Matriks Konfusi dan Hasil Klasifikasi Utama

Tabel 7 menyajikan nilai Matriks Konfusi untuk kedua sistem eksperimental pada set uji 50 ulasan. Tabel 8 menunjukkan metrik performa yang diturunkan yang dihitung menggunakan persamaan evaluasi.

Tabel 7. Matriks Konfusi - Set Uji (50 ulasan)

Sistem	TP	TN	FP	FN
Sistem A - Dengan Normalisasi BMM (Diusulkan)	24	18	3	5
Sistem B - Tanpa Normalisasi (Baseline)	19	14	7	10

Tabel 8. Metrik Performa - Terverifikasi Secara Matematis dari Tabel 7

Sistem	Akurasi	Presisi	Recall	F1-Score
Sistem A - Dengan Normalisasi BMM	84,00 %	88,89 %	82,76 %	85,71 %
Sistem B - Tanpa Normalisasi	66,00 %	73,08 %	65,52 %	69,09 %
Peningkatan (Atas B)	+18,00 pp	+15,81 pp	+17,24 pp	+16,62 pp

Peningkatan akurasi 18 poin persentase adalah ukuran paling langsung dari kontribusi modul normalisasi dalam praktik. Melihat Matriks Konfusi, perbedaan paling jelas ada pada False Negative: Sistem B memiliki FN=10 (melewatkan 10 ulasan positif), sementara Sistem A hanya FN=5 - pengurangan 50%. Hal ini masuk akal karena banyak ulasan positif dalam BMM mengandung intensifier seperti "pe gaga" atau "mantap skali" yang tidak dapat dibaca oleh Sistem B. Presisi tinggi 88,89% pada Sistem A berarti ketika sistem memprediksi ulasan positif, prediksi tersebut benar hampir 9 dari 10 kali - tingkat keandalan yang praktis bagi pemilik usaha.

5. 10-Fold Cross Validation

Untuk memverifikasi stabilitas hasil, 10-Fold Cross Validation dijalankan pada seluruh dataset 250 ulasan. Setiap fold menggunakan 25 ulasan berbeda sebagai set uji,

sehingga setiap ulasan dievaluasi tepat satu kali.

Tabel 9. Hasil 10-Fold Cross Validation - Akurasi per Fold

Fold	Sistem A (Dengan Normalisasi BMM)	Sistem B (Tanpa Normalisasi)
Fold 1	80,00%	62,00%
Fold 2	84,00%	68,00%
Fold 3	76,00%	64,00%
Fold 4	84,00%	64,00%
Fold 5	80,00%	66,00%
Fold 6	88,00%	72,00%
Fold 7	84,00%	66,00%
Fold 8	80,00%	68,00%
Fold 9	84,00%	64,00%
Fold 10	80,00%	66,00%
Rata-rata ± Std Dev	82,00% ± 3,22%	66,00% ± 3,22%

Hasil cross validation konsisten dengan temuan set uji. Sistem A mempertahankan rata-rata akurasi 82,00% ± 3,22% di semua sepuluh fold, sementara Sistem B rata-rata 66,00% ± 3,22%. Standar deviasi rendah (3,22%) untuk kedua sistem menunjukkan bahwa hasilnya tidak bergantung pada ulasan mana yang masuk ke set uji - perbedaan performa antara kedua sistem didorong secara struktural oleh modul normalisasi.

6. Analisis Komparatif dengan Studi Terkait

Tabel 10. Analisis Komparatif Lengkap: Penelitian Ini vs. Studi Konteks Manado Sebelumnya

Penulis (Tahun)	Domain	Algoritma	Akurasi	F1-Score	Norm. BMM?
Petronela (2023)	Hotel / Manado	Naive Bayes	76,20%	T/L	Tidak
Wongkar (2020)	Twitter / Manado	Naive Bayes	75,58%	T/L	Tidak
Haris (2025)	Aplikasi UMKM	NBC + TF-IDF	65,00%	T/L	Tidak
Penelitian Ini	Furnitur /	NBC + Lek	84,00%	85,71%	Ya (120

(2026)	Manado	sikon BMM	entri)

(T/L = Tidak dilaporkan dalam studi asli) Sistem yang diusulkan mencapai akurasi tertinggi di antara semua studi konteks Manado yang ditinjau, dengan margin +7,80 poin persentase atas hasil terbaik sebelumnya (Petronela, 2023 pada 76,20%) dan +19,00 pp atas Haris (2025). Yang paling penting, ini adalah satu-satunya studi dalam kelompok perbandingan yang secara eksplisit menangani masalah dialek BMM.

7. Analisis Kesalahan

Dari 50 ulasan uji, 8 salah diklasifikasikan oleh Sistem A: 3 False Positive (diprediksi positif, sebenarnya negatif) dan 5 False Negative (diprediksi negatif, sebenarnya positif). Dua pola kesalahan berulang muncul: (1) Ulasan bersentimen campuran yang menggabungkan evaluasi produk positif dengan keluhan layanan atau logistik negatif dalam teks yang sama - sinyal positif menghasilkan bukti bag-of-words yang lebih kuat dari sinyal negatif; (2) Polaritas kata yang bergantung konteks: kata seperti "murah" bisa positif (harga bagus) atau negatif (kualitas rendah) tergantung konteks - perbedaan yang tidak dapat diandalkan oleh model bag-of-words.

8. Ringkasan Modul Sistem

Tabel 11. Ringkasan Modul Sistem

Modul	Aktor	Fungsi	Teknologi
Formulir Pengiriman Ulasan	Pelanggan	Kirim ulasan via web/QR; memicu pipeline preprocessing secara otomatis	Laravel Blade + AJAX
Login Admin	Admin/Pemilik	Autentikasi aman ke dashboard	Laravel Auth + Bcrypt
Dashboard Sentimen	Admin/Pemilik	Grafik donat, tren bulanan, word cloud negatif	Chart.js + Laravel
Manajer Ulasan	Admin/Pemilik	Jelajah, filter, ekspor ulasan dengan label	Laravel DataTables + CSV
Manajer Leksikon	Admin/Pemilik	Tambah/edit/hapus pasangan kata BMM (antarmuka CRUD)	Laravel Resource Controller

API Klasifikasi ML	Sistem (internal)	Menerima teks, mengembalikan label + skor kepercayaan	Python + Flask + Pickle

D. Kesimpulan dan Saran

1. Kesimpulan

Penelitian ini membangun dan mengevaluasi sistem analisis sentimen berbasis web yang berfungsi untuk CV Talongka jaya, sebuah UMKM furnitur di Manado, Indonesia. Kontribusi utama adalah modul normalisasi leksikon dialek Bahasa Melayu Manado (BMM) kustom - 120 pasang kata yang memetakan kosakata BMM ke Bahasa Indonesia standar, diterapkan sebagai langkah preprocessing sebelum classifier Multinomial Naive Bayes. Sistem mencapai akurasi 84,00%, presisi 88,89%, recall 82,76%, dan F1-Score 85,71% pada set uji 50 ulasan, dibandingkan akurasi 66,00% dan F1-Score 69,09% untuk baseline tanpa normalisasi - peningkatan akurasi 18 poin persentase.

Ablation study mengonfirmasi bahwa langkah normalisasi berkontribusi pada peningkatan performa terbesar di antara semua tahap preprocessing (+10,00 pp akurasi, Konfigurasi C ke D). Rata-rata 10-Fold Cross Validation sebesar 82,00% ± 3,22% versus 66,00% ± 3,22% untuk baseline mengonfirmasi bahwa perbedaan performa stabil dan tidak bergantung pada pembagian acak tertentu.

Hasil ini membangun benchmark akurasi baru untuk analisis sentimen Naive Bayes dalam konteks linguistik Manado, melampaui ketiga studi konteks Manado sebelumnya. Yang lebih luas, penelitian ini menunjukkan bahwa leksikon dialek berbasis aturan yang sederhana dapat menutup kesenjangan performa yang signifikan untuk NLP bahasa daerah - tanpa memerlukan model neural besar, korpus pelatihan beranotasi dialek, atau sumber daya komputasi GPU.

2. Saran

Untuk peneliti masa depan, langkah berikutnya yang paling berdampak adalah analisis sentimen berbasis aspek - mengklasifikasikan opini pada dimensi produk tertentu (keahlian, pengiriman, layanan, harga) secara terpisah daripada memaksakan satu label biner. Leksikon BMM harus dipublikasikan sebagai sumber daya open-source di

GitHub atau Zenodo untuk mendukung komunitas NLP bahasa daerah Indonesia yang lebih luas. Untuk praktisi, mengintegrasikan pengumpulan ulasan dengan WhatsApp Business API akan memungkinkan akuisisi data yang sepenuhnya otomatis. Untuk peneliti yang menginginkan akurasi lebih tinggi, fine-tuning IndoBERT pada teks yang dinormalisasi BMM adalah langkah alami berikutnya yang dapat mendorong performa jauh melampaui apa yang dapat dicapai model Naive Bayes bag-of-words.

Daftar Pustaka

Haris, A. (2025). Penerapan Algoritma Naive Bayes Untuk Klasifikasi Sentimen Ulasan Pengguna Aplikasi Pidi UMKM. Jurnal Teknologi Informasi dan Sistem Informasi, 5(1), 45-58.

IJCAONLINE. (2023). Systematic review on text normalization techniques and its approach to non-standard words. International Journal of Computer Applications, 185(33). https://doi.org/10.5120/ijca2023923106

Jumawan, J., Soesanto, E., Cahya, F., Putri, C. A., Permatasari, S. A., Setyakinasti, S., & Ottay, M. L. (2024). Pengaruh Online Customer Reviews terhadap Keputusan Pembelian di Marketplace Shopee. SENTRI: Jurnal Riset Ilmiah, 3(6), 2854-2862.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159-174. https://doi.org/10.2307/2529310

Liu, B. (2020). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781108639286

McCallum, A., & Nigam, K. (1998). A comparison of event models for Naive Bayes text classification. AAAI-98 Workshop on Learning for Text Categorization, 752, 41-48.

Naseem, U., Eklund, P., Razzak, I., & Musial, K. (2024). Machine learning based framework for fine-grained word segmentation and enhanced text normalization for low resourced language. PLOS ONE, 19(12), e0315237.

https://doi.org/10.1371/journal.pone.0315237

Nguyen, T. L., Le, H. T., & Nguyen, T. M. (2025). ViDia2Std: A parallel corpus and methods for low-resource Vietnamese dialect-to-standard translation. ArXiv Preprint. https://arxiv.org/abs/2603.10211

Petronela, J. (2023). Sentiment analysis of visitor reviews on star hotels in Manado City. International Journal of Computer and Information System (IJCIS), 4(2), 67-74. https://doi.org/10.29040/ijcis.v4i2.101

Ridho, M. R., Arnomo, S. A., Fajrah, N., & Fifi, F. (2023). Analisis Sentimen Ulasan Penjualan UMKM Sosial Melalui Media. Jurnal Desain Dan Analisis Teknologi, 2(2), 169-175. https://doi.org/10.58520/jddat.v2i2.35

Setiadi, D. R. I. M., Marutho, D., & Setiyanto, N. A. (2024). Comprehensive exploration of machine and deep learning classification methods for aspect-based sentiment analysis. Journal of Future Artificial Intelligence and Technology, 1(1), 12-22. https://doi.org/10.62411/faith.2024-3

Sneddon, J. N. (1983). Diglossia and Manado Malay. NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia, 17, 1-17.

Walasary, T. (2022). Survey paper tentang analisis sentimen. Jurnal Konstelasi, 1(1), 12-19.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 1, 29-39.

Wongkar, M. (2020). Sentiment analysis using Naive Bayes algorithm of the data crawler: Twitter. Journal of Information Technology and Computer Science, 2(1), 1-8.