

## Penerapan Naive Bayes untuk Klasifikasi E-Lapor Kabupaten Sumedang

Muhamad Fadlan Asmawi<sup>1,\*</sup>, Agun Guntara<sup>2</sup>

<sup>1</sup> Informatika; Universitas Sebelas April; Sumedang, telp/fax : 082131867869; e-mail: [220660121059@student.usap.ac.id](mailto:220660121059@student.usap.ac.id).

<sup>2</sup> Sistem Informasi; Universitas Sebelas April; Sumedang, 08987922009; e-mail: [agun@unsap.ac.id](mailto:agun@unsap.ac.id)

\* Korespondensi: e-mail: [fadlanasmawi7@gmail.com](mailto:fadlanasmawi7@gmail.com)

Diterima: 25 November 2025; Review: 06 Desember 2025; Disetujui: 18 Desember 2025

Cara sitasi: Asmawi, M. Fadlan, Guntara, Agun. 2025. Penerapan Naive Bayes untuk Klasifikasi E-Lapor Kabupaten Sumedang. Information Management for Educators and Professionals. Vol 10 (2): 143-152

**Abstrak:** Sistem E-Lapor (SP4N-LAPOR!) di Kabupaten Sumedang menghadapi tantangan dalam mengklasifikasikan volume laporan masyarakat yang masuk, di mana proses disposisi manual ke instansi terkait berjalan lambat dan rentan terhadap kesalahan. Penelitian ini bertujuan untuk mengimplementasikan algoritma *Naive Bayes Classifier* (NBC) dengan pembobotan TF-IDF untuk mengotomatisasi proses klasifikasi tersebut. Metodologi penelitian menerapkan tahapan pra-pemrosesan teks (*case folding, tokenizing, stopword removal*) tanpa *stemming*, diikuti dengan pembagian data 80:20 dan evaluasi menggunakan *confusion matrix*. Hasil penelitian menunjukkan kinerja model yang sangat rendah akibat kendala data yang tidak seimbang (*imbalanced*), dengan akurasi akhir hanya 24,14%, *Weighted F1-Score* 10%, dan Presisi 6%. Temuan utama dari penelitian ini adalah bahwa kegagalan kinerja bukan disebabkan oleh kelemahan algoritma NBC, melainkan oleh masalah fundamental pada dataset yang digunakan: jumlah data yang sangat tidak memadai (*data sparsity*) dengan total hanya ~145 sampel bersih, serta distribusi kelas yang sangat timpang. Temuan ini menyimpulkan bahwa dataset E-Lapor Sumedang saat ini tidak layak untuk pelatihan model otomatis tanpa penambahan data dan penerapan teknik penyeimbangan data.

**Kata kunci:** E-Lapor, Klasifikasi Teks, Naive Bayes, SP4N-LAPOR!, Data Tidak Seimbang

**Abstract:** The E-Lapor (SP4N-LAPOR!) system in Sumedang Regency faces challenges in classifying incoming public reports, where manual disposition is slow and error-prone. This research aims to implement the Naive Bayes Classifier (NBC) algorithm with TF-IDF weighting to automate this classification. The methodology applies text pre-processing stages (*case folding, tokenizing, stopword removal*) without *stemming*, followed by an 80:20 data split and evaluation using a *confusion matrix*. The results showed extremely low model performance due to imbalanced data constraints, with a final accuracy of only 24.14%, *Weighted F1-Score* of 10%, and *Precision* of 6%. The main finding is that this failure was not caused by NBC algorithm weaknesses, but by fundamental issues within the dataset: severe data sparsity (only ~145 clean samples) and a highly skewed class distribution. It is concluded that the current E-Lapor Sumedang dataset is not feasible for automated model training without data augmentation and balancing techniques.

**Keywords:** E-Lapor, Text Classification, Naive Bayes, SP4N-LAPOR!, Imbalanced Data

## 1. Pendahuluan

Pelayanan publik merupakan elemen fundamental dalam tata kelola pemerintahan yang baik di era digital. Seiring dengan kemajuan teknologi, ekspektasi masyarakat terhadap kecepatan dan transparansi respons pemerintah semakin meningkat. Pemerintah Indonesia merespons hal ini melalui Sistem Pengelolaan Pengaduan Pelayanan Publik Nasional (SP4N-LAPOR!) sebagai kanal terpadu aspirasi masyarakat [1], [2]. Di Kabupaten Sumedang, pemanfaatan E-Lapor telah menghasilkan volume data laporan harian yang signifikan.

Namun, besarnya volume data teks tidak terstruktur ini menciptakan tantangan operasional serius. Saat ini, proses klasifikasi laporan masih dilakukan secara manual. Petugas admin harus membaca satu per satu isi laporan untuk memverifikasi dan mendisposisikannya ke dinas terkait [3]. Proses manual ini tidak efisien, memakan waktu, dan rentan terhadap *human error* atau subjektivitas, yang pada akhirnya memperlambat *response time* pemerintah. Oleh karena itu, dibutuhkan otomatisasi menggunakan pendekatan *Text Mining* dan *Machine Learning* untuk mempercepat proses disposisi tersebut [4].

Salah satu algoritma klasifikasi teks yang terbukti andal, cepat, dan efisien secara komputasi adalah *Naive Bayes Classifier* (NBC) [5], [6]. Meskipun terdapat metode lain seperti *Support Vector Machine* (SVM) atau *Deep Learning*, Naive Bayes dipilih dalam penelitian ini karena kesederhanaannya yang efektif untuk klasifikasi teks dasar dengan sumber daya komputasi yang ringan [7], [8].

Novelty dari penelitian ini terletak pada fokus implementasinya. Jika mayoritas penelitian SP4N-LAPOR! sebelumnya lebih menitikberatkan pada analisis sentimen positif/negatif untuk mengukur kepuasan publik [5], [9], [10], penelitian ini secara spesifik berfokus pada klasifikasi kategori instansi disposisi otomatis di tingkat pemerintah daerah Kabupaten Sumedang. Hal ini mengisi celah penelitian (*research gap*) mengenai kelayakan teknis otomatisasi disposisi laporan di lingkup pemerintah kabupaten yang seringkali memiliki karakteristik data unik dan terbatas (*data sparsity*).

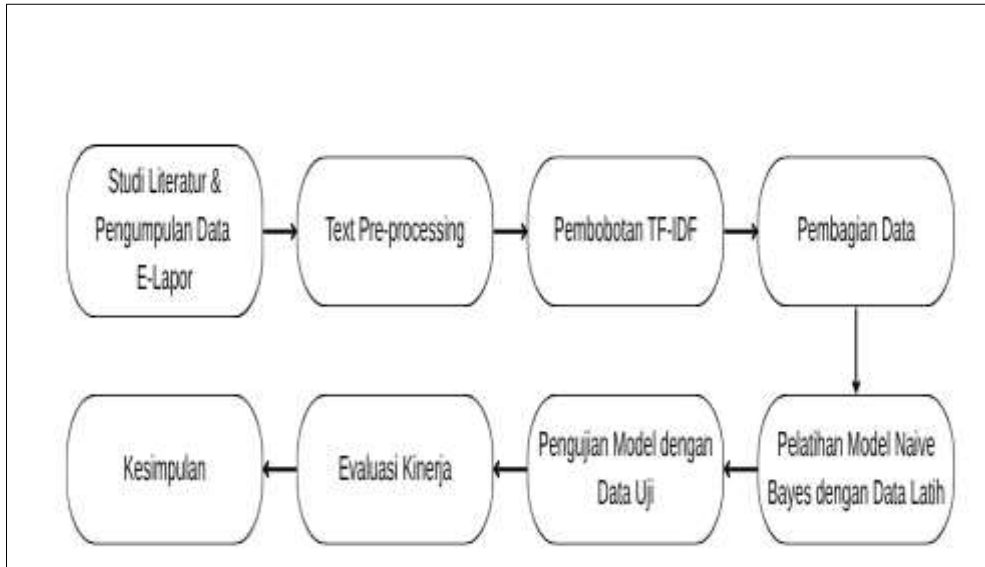
Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah, Bagaimana menerapkan algoritma Naive Bayes untuk mengklasifikasikan laporan masyarakat secara otomatis dan sejauh mana kelayakan kinerja model tersebut jika diterapkan pada data E-Lapor Kabupaten Sumedang.

Agar penelitian lebih terarah, ruang lingkup penelitian dibatasi pada: (1) Objek penelitian adalah data teks laporan masyarakat yang masuk ke admin E-Lapor Kabupaten Sumedang periode tahun 2024; (2) Data yang diproses hanya berupa teks aduan utama (tidak termasuk lampiran gambar/video); dan (3) Klasifikasi dilakukan pada kategori Organisasi Perangkat Daerah (OPD) atau dinas tujuan.

Tujuan penelitian ini adalah membangun model klasifikasi otomatis menggunakan Naive Bayes dan mengevaluasi kelayakannya. Penilaian kelayakan tidak hanya didasarkan pada Akurasi semata, tetapi juga menggunakan metrik *Precision*, *Recall*, dan *F1-Score* untuk memastikan model mampu menangani ketidakseimbangan kelas *imbalanced data* yang umum terjadi pada data pengaduan masyarakat [11].

## 2. Metode Penelitian

Penelitian ini mengadopsi pendekatan kuantitatif yang berfokus pada *Text Mining* dan *Machine Learning*. Proses penelitian melibatkan serangkaian tahapan sistematis, dimulai dari pengumpulan data, studi literatur, pembersihan data (pra-pemrosesan), pembobotan fitur, implementasi algoritma, hingga evaluasi akhir model.



Sumber: Hasil Penelitian (2025)

Gambar 1. Kerangka Kerja Penelitian

### Sumber Data

Sumber data dalam penelitian ini bersifat sekunder, yakni arsip Laporan Ringkas E-Lapor (SP4N-LAPOR!) Kabupaten Sumedang selama periode 2024. Data awal diperoleh dalam format tabular (*spreadsheet*) yang mencakup beberapa atribut utama seperti ID Laporan, Tanggal, Isi Laporan, dan Instansi Tujuan. Sebagai contoh format laporan, satu entri data memuat narasi teks keluhan masyarakat pada kolom "Isi Laporan" dan nama dinas terkait pada kolom "Instansi Tujuan". Dari dataset tersebut, fokus penelitian tertuju pada dua kolom esensial: a). Isi Laporan: Digunakan sebagai variabel fitur (X), yang berisi teks aduan masyarakat. b). Instansi Terdisposisi (atau Kategori): Digunakan sebagai variabel label/target (Y), yang menentukan klasifikasi (misal: "Dinas PUPR", "Dinas Kesehatan").

Mengingat distribusi kategori instansi yang sangat bervariasi dan cenderung tidak seimbang (*imbalanced*) pada variabel target (Y) [11], proses pemisahan data dilakukan menggunakan teknik *Stratified Random Sampling*. Teknik stratifikasi ini diterapkan untuk memastikan proporsi setiap kategori instansi (kelas) tetap terjaga keseimbangannya baik pada data latih maupun data uji. Selain itu, parameter *random\_state* diset pada nilai 42 untuk menjamin konsistensi dan reproduktibilitas hasil pembagian data.

### Tahapan Text Pre-processing

Algoritma *machine learning* tidak dapat memproses teks mentah secara langsung. Teks aduan mentah masih sarat akan *noise*, seperti variasi huruf kapital, tanda baca, dan kata-kata umum yang tidak relevan (kata henti) [12]. Untuk membersihkan dan menstandarisasi data, diterapkan serangkaian tahapan *Text Pre-processing*, yang dalam berbagai studi terbukti vital untuk meningkatkan performa model [11]. Tahapan tersebut meliputi: Case Folding: Menyeragamkan seluruh teks menjadi format huruf kecil (*lowercase*). Tokenizing: Memecah struktur kalimat menjadi unit-unit kata (token) individual. Stopword Removal: Memfilter dan membuang kata-kata henti (kata umum tanpa makna khusus, seperti "yang", "di", "dari", "ini") menggunakan kamus *stopwords* Bahasa Indonesia [4].

Pentingnya tahapan ini tidak dapat dipandang sebelah mata, terutama mengingat karakteristik morfologi Bahasa Indonesia yang kaya akan imbuhan (afiks). Meskipun penelitian ini membatasi diri tanpa menggunakan *stemming* karena keterbatasan sumber daya, proses *Stopword Removal* menjadi filter pertahanan utama untuk mereduksi *noise*. Kata-kata umum seperti "yang", "dan", atau "di" memiliki frekuensi kemunculan yang sangat tinggi dalam hampir semua dokumen. Jika tidak dibuang, kata-kata ini akan mendominasi perhitungan frekuensi dan mengaburkan kata kunci substantif (seperti "jalan", "KTP", "banjir") yang sebenarnya membawa

informasi pembeda antar kategori dinas [12]. Penghapusan *stopwords* secara efektif mengurangi dimensi fitur input, sehingga algoritma dapat lebih fokus pada terminologi inti yang relevan.

### Pembobotan Kata (TF-IDF)

Pasca tahapan pra-pemrosesan, data teks harus ditransformasikan menjadi representasi numerik agar dapat dikalkulasi oleh algoritma *machine learning*. Penelitian ini menerapkan metode Term Frequency-Inverse Document Frequency (TF-IDF) sebagai teknik pembobotan fitur [13]. Metode ini bekerja dengan memberikan bobot signifikansi yang lebih tinggi pada kata-kata yang sering muncul dalam satu dokumen spesifik (laporan) namun jarang ditemukan dalam keseluruhan koleksi dokumen lain. Hal ini membuat kata tersebut dianggap sebagai fitur unik yang penting untuk membedakan antar kategori.

Term Frequency (TF) menghitung frekuensi kemunculan sebuah kata (*term*) di dalam satu dokumen tertentu:

$$tf(t, d) = \frac{\text{jumlah kemunculan term } t \text{ dalam dokumen } d}{\text{total term dalam dokumen } d} \quad (1)$$

Inverse Document Frequency (IDF) mengukur seberapa umum atau jarang sebuah kata muncul di seluruh koleksi dokumen, yang dihitung dengan logaritma dari total dokumen dibagi jumlah dokumen yang mengandung kata tersebut:

$$idf(t, N) = \log \left( \frac{N}{df_t} \right) \quad (2)$$

Bobot akhir TF-IDF untuk sebuah kata adalah hasil perkalian antara nilai TF dan IDF:

$$w(t, d) = tf(t, d) \times idf(t, N) \quad (3)$$

Secara konseptual, mekanisme pembobotan ini memegang peranan vital dalam membedakan konteks laporan. Dalam korpus data pengaduan masyarakat, kata-kata seperti 'mohon', 'bantu', atau 'lapor' cenderung muncul di hampir semua kategori dinas, sehingga nilai IDF-nya akan sangat rendah (mendekati nol). Sebaliknya, kata spesifik seperti 'e-KTP' (yang hanya relevan untuk Disdukcapil) atau 'aspal' (untuk Dinas PUPR) akan memiliki nilai IDF yang tinggi. Dengan demikian, algoritma Naive Bayes tidak akan tertipu oleh kata-kata umum yang memiliki frekuensi tinggi, melainkan akan fokus pada kata-kata berbobot tinggi yang menjadi *signature* atau penanda unik dari masing-masing instansi tujuan [13]. Tanpa pembobotan ini, model akan kesulitan membedakan prioritas fitur dalam teks yang pendek.

### Algoritma Klasifikasi Naive Bayes

Pilar utama klasifikasi dalam penelitian ini adalah Naive Bayes Classifier (NBC). Algoritma ini dipilih berdasarkan rekam jejaknya yang terbukti andal untuk klasifikasi teks, efisiensi komputasi yang tinggi, dan kecepatan pemrosesan [14], [11]. NBC beroperasi dengan mengaplikasikan Teorema Bayes, yang menghitung probabilitas suatu dokumen (laporan) termasuk dalam suatu kelas (kategori) berdasarkan data latih.

$$P(C | X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (4)$$

Keterangan:

$P(C | X)$ : Probabilitas posterior (peluang kelas  $C$  benar setelah data  $X$  diobservasi).

$P(X | C)$ : Probabilitas likelihood (peluang data  $X$  muncul jika diketahui kelas  $C$ ).

$P(C)$ : Probabilitas prior (probabilitas awal kemunculan kelas  $C$ ).

$P(X)$ : Probabilitas marginal data  $X$ .

Dalam konteks klasifikasi teks, penelitian ini secara spesifik menggunakan varian Multinomial Naive Bayes (MNB), yang sangat efektif untuk data fitur diskrit seperti

frekuensi kata (TF-IDF) [6]. Rumus MNB digunakan untuk mencari probabilitas tertinggi dari sebuah dokumen  $D$  untuk setiap kelas  $C$  yang ada:

$$P(C | D) \propto P(C) \prod_{i=1}^n P(t_i | C) \tag{5}$$

Konfigurasi model dan parameter untuk memastikan reproduibilitas hasil, implementasi model dalam penelitian ini menggunakan spesifikasi parameter sebagai berikut:

Algoritma: Multinomial Naive Bayes. Parameter Smoothing: Nilai  $\alpha = 1.0$  (Laplace Smoothing) diterapkan untuk mencegah probabilitas nol pada kata yang tidak muncul dalam data latih. Dataset: Model dilatih menggunakan total 145 sampel data bersih yang terdistribusi ke dalam 16 kategori instansi. Validasi: Pengujian dilakukan menggunakan skema *split* 80:20 dengan pengaturan `random_state = 42` untuk menjaga konsistensi pengacakan data.

Model akan memprediksi laporan tersebut masuk ke dalam kategori dengan nilai probabilitas posterior  $P(C | D)$  yang paling tinggi.

### Skenario Pengujian dan Metrik Evaluasi

Untuk memvalidasi keandalan model yang telah dibangun, dataset yang telah melalui proses pembobotan TF-IDF dipartisi menjadi dua subset menggunakan teknik *Stratified Random Sampling* dengan rasio 80:20. Dalam skema ini, 80% data dialokasikan sebagai data latih untuk proses pembelajaran model, sedangkan 20% sisanya digunakan sebagai data uji untuk tahap validasi. Penerapan teknik stratifikasi ini sangat krusial untuk menjaga konsistensi proporsi kelas antara data latih dan data uji, mengingat karakteristik dataset yang digunakan tidak seimbang.

Evaluasi kinerja klasifikasi dilakukan menggunakan metode *Confusion Matrix* [15]. Metode ini menyajikan pemetaan komprehensif antara hasil prediksi yang diberikan oleh model dengan label kategori aktual pada data uji.

Tabel 1. Struktur Konseptual Confusion Matrix (Biner)

	Prediksi: Positif	Prediksi: Negatif
Aktual: Positif	True Positive (TP)	False Negative (FN)
Aktual: Negatif	False Positive (FP)	True Negative (TN)

Sumber: Hasil Penelitian (2025)

Berdasarkan hasil pemetaan pada *Confusion Matrix* tersebut, performa model diukur menggunakan tiga metrik evaluasi standar [15], yaitu: Akurasi (*Accuracy*): Mengukur tingkat ketepatan model secara keseluruhan dalam memprediksi seluruh kelas data dengan benar. Presisi (*Precision*): Mengukur tingkat keakuratan model dalam menetapkan prediksi positif, atau rasio jumlah prediksi benar positif dibandingkan dengan total data yang diprediksi positif. Recall (Sensitivitas): Mengukur kemampuan model dalam mengenali kembali seluruh data yang sebenarnya bernilai positif dari total data relevan yang ada.

$$Akurasi = (TP + TN) / (TP + TN + FP + FN)$$

$$Presisi = TP / (TP + FP)$$

$$Recall = TP / (TP + FN)$$

### 3. Hasil dan Pembahasan

Implementasi algoritma Naive Bayes pada dataset E-Lapor Sumedang telah berhasil dieksekusi. Namun, hasil pengujian pada 20% data uji menunjukkan kinerja model yang sangat rendah. Hal ini tidak hanya terlihat dari nilai Akurasi global yang hanya mencapai 24,14%, tetapi juga dikonfirmasi oleh metrik evaluasi lainnya yang sangat minim. Nilai rata-rata tertimbang (*weighted average*) untuk F1-Score tercatat hanya sebesar 10%, dengan tingkat Presisi sebesar 6%. Temuan statistik ini mengindikasikan secara kuat bahwa konfigurasi model saat ini gagal mengenali pola data dan tidak mampu memprediksi mayoritas kategori dengan benar. Rincian kinerja model untuk setiap kategori instansi disajikan secara lengkap pada Tabel 2.

Tabel 2. Laporan Klasifikasi Rinci Model Naive Bayes

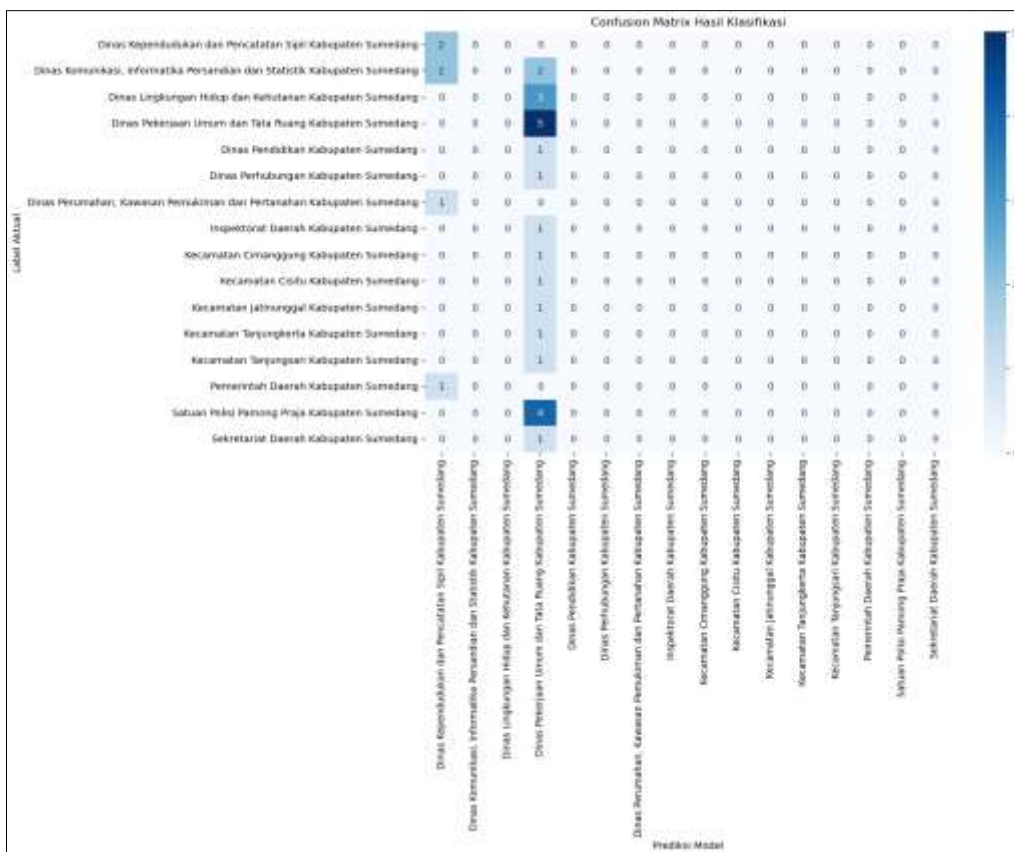
Kategori (Instansi)	Precision	Recall	F1-Score	Support
Disdukcapil	0.33	1.00	0.50	2
Diskominfo	0.00	0.00	0.00	4

DLHK	0.00	0.00	0.00	3
PUPR	0.22	1.00	0.36	5
Disdik	0.00	0.00	0.00	1
Dishub	0.00	0.00	0.00	1
Disperkimtan	0.00	0.00	0.00	1
Inspektorat	0.00	0.00	0.00	1
Kec. Cimanggung	0.00	0.00	0.00	1
Kec. Cisu	0.00	0.00	0.00	1
Kec. Jatinunggal	0.00	0.00	0.00	1
Kec. Tanjungkerta	0.00	0.00	0.00	1
Kec. Tanjungsari	0.00	0.00	0.00	1
Pemda	0.00	0.00	0.00	1
Satpol PP	0.00	0.00	0.00	4
Setda	0.00	0.00	0.00	1
<b>Accuracy</b>			<b>0.24</b>	<b>29</b>
<b>Macro Avg</b>	<b>0.03</b>	<b>0.12</b>	<b>0.05</b>	<b>29</b>
<b>Weighted Avg</b>	<b>0.06</b>	<b>0.24</b>	<b>0.10</b>	<b>29</b>

Sumber: Hasil Penelitian (2025)

### Analisis Kinerja Model

Capaian akurasi sebesar 24,14% (dengan *F1-score* rata-rata tertimbang hanya 10%) memberikan indikasi kuat bahwa model gagal melakukan generalisasi pola yang bermakna dari data latih. Berdasarkan bedah data pada Tabel 2 dan visualisasi *Confusion Matrix* (Gambar 2), kegagalan ini bermanifestasi dalam dua fenomena anomali utama.



Sumber: Hasil Penelitian (2025)

Gambar 2. Confusion Matrix Hasil Klasifikasi

Pertama, ketidakmampuan deteksi pada kelas minoritas. Terlihat adanya pola *zero-prediction* yang ekstrem, di mana 14 dari total 16 kategori instansi mencatatkan nilai presisi, *recall*, dan *f1-score* sebesar 0,00. Angka ini merefleksikan bahwa model mengalami kesulitan fatal dalam mengenali fitur unik pada kategori-kategori tersebut, sehingga model cenderung sama

sekali tidak memprediksi atau mengabaikan keberadaan kelas-kelas minoritas ini dalam hasil akhirnya.

Kedua, bias prediksi terhadap kelas mayoritas. Model menunjukkan fenomena yang menyerupai "ilusi akurasi" pada kategori dominan seperti Dinas PUPR dan Disdukcapil. Pada kategori ini, nilai *Recall* mencapai angka sempurna (1.00), namun nilai *Precision*-nya sangat rendah (contoh: Dinas PUPR hanya 0,22). Secara statistik, ini menunjukkan bahwa model melakukan prediksi secara agresif (menebak "Dinas PUPR") pada hampir semua data yang masuk. Akibatnya, meskipun seluruh laporan asli Dinas PUPR berhasil terdeteksi (*True Positive*), strategi ini menghasilkan banyak kesalahan identifikasi (*False Positive*) karena laporan dari dinas lain juga turut diklaim sebagai laporan PUPR. Hal ini terkonfirmasi pada *Confusion Matrix* (Gambar 2), di mana sebaran prediksi menumpuk secara vertikal pada kolom kategori mayoritas.

### Akar Permasalahan: *Data Sparsity* dan Ketimpangan Distribusi Kelas

Rendahnya performa klasifikasi model (Akurasi: 24,14%) bukanlah indikasi dari kegagalan algoritma Naive Bayes secara isolatif, melainkan manifestasi dari karakteristik dataset yang tidak ideal untuk pembelajaran mesin. Berdasarkan analisis mendalam terhadap pola prediksi dan perbandingan dengan literatur sejenis, akar permasalahan dapat dipetakan menjadi dua faktor fundamental:

#### Kuantitas Data yang Tidak Memadai (*Data Sparsity*)

Faktor penghambat utama adalah *data sparsity*. Seperti terlihat pada kolom *support* (Tabel 2), total data uji hanya berjumlah 29 laporan, yang mengimplikasikan bahwa total dataset bersih (*clean dataset*) yang tersedia untuk pembelajaran hanya berkisar 145 sampel.

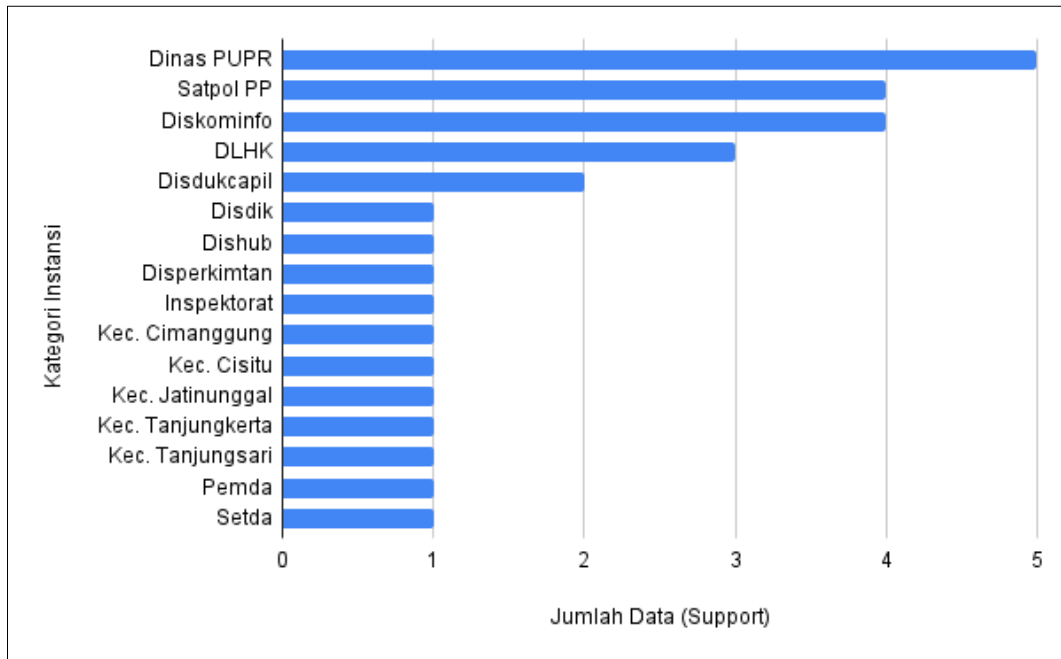
Secara analitis, jumlah ini sangat tidak memadai jika disandingkan dengan studi serupa yang berhasil. Sebagai komparasi, penelitian Rozi et al. [3] yang juga melakukan klasifikasi teks pada platform SP4N-LAPOR! mampu mencapai performa yang baik karena didukung oleh dataset berskala ribuan. Pada penelitian ini, keterbatasan sampel menyebabkan model gagal menangkap variabilitas linguistik yang kompleks dari 16 kategori berbeda. Tidak ada algoritma *supervised learning* yang dapat melakukan generalisasi yang valid hanya dengan rata-rata ~9 sampel per kelas.

Dampak dari *data sparsity* ini teramplifikasi secara eksponensial akibat absennya tahapan *stemming*. Dalam linguistik komputasi, *stemming* berfungsi untuk memetakan varian morfologis ke bentuk dasar yang sama (contoh: 'memperbaiki', 'perbaikan', 'diperbaiki' menjadi 'baik'). Tanpa proses ini, model Naive Bayes memperlakukan ketiga kata tersebut sebagai entitas fitur yang sepenuhnya terpisah dan tidak berhubungan (orthographically independent). Akibatnya, bobot probabilitas yang seharusnya terkumpul pada satu konsep 'perbaikan', malah terpecah (*diluted*) ke dalam banyak fitur dengan frekuensi yang sangat kecil [11]. Dalam kondisi dataset yang hanya berjumlah 145 sampel, pemecahan bobot ini sangat fatal karena membuat sinyal pola menjadi terlalu lemah untuk dideteksi oleh model, yang pada akhirnya berkontribusi langsung pada rendahnya nilai *Recall* pada kategori-kategori spesifik.

#### Data yang Sangat Tidak Seimbang

Terlihat jelas adanya dominasi kategori mayoritas (seperti Dinas PUPR) dibandingkan kategori minoritas yang hanya memiliki 1-2 sampel. Kondisi ini menyebabkan performa Naive Bayes runtuh karena ketergantungan algoritma ini pada Probabilitas Prior. Sesuai teorema Bayes:  $P(C|X) \propto P(X|C) \times P(C)$ , dimana:  $P(X|C)$  adalah *likelihood* (kemiripan kata-kata dalam laporan dengan kategori).  $P(C)$  adalah *prior probability* (seberapa sering kategori muncul di seluruh dataset). Faktor ini menjadi penyebab paling kritis dari bias prediksi model ("collapse"). Distribusi kelas yang sangat timpang pada dataset E-Lapor Sumedang disajikan pada Gambar 3.

Pada kasus kategori minoritas (misal: Dinas Pendidikan), nilai  $P(C)$  sangat kecil (mendekati nol). Akibatnya, meskipun sebuah teks laporan memiliki kata-kata yang sangat relevan dengan pendidikan (nilai *likelihood* tinggi), hasil perkalian akhirnya akan tetap kalah besar dibandingkan kategori mayoritas yang memiliki nilai  $P(C)$  sangat dominan.



Sumber: Hasil Penelitian (2025)

Gambar 3. Grafik Distribusi Kelas pada Dataset E-Lapor Sumedang

Secara matematis, model belajar bahwa memprediksi kategori minoritas memiliki risiko kesalahan yang besar. Hal ini menjelaskan mengapa akurasi pada penelitian ini (24.14%) tertinggal jauh dibandingkan penelitian Rahmawati [5] yang menerapkan Naive Bayes pada dataset ulasan yang lebih seimbang. Kegagalan ini mengonfirmasi bahwa tanpa teknik penyeimbang data (*resampling*) seperti SMOTE, Naive Bayes akan cenderung bias ke kelas mayoritas.

Implikasi dari ketimpangan distribusi ini tidak hanya berdampak pada metrik akurasi, tetapi juga pada validitas operasional model jika diterapkan di dunia nyata. Dengan bias yang begitu kuat terhadap kelas mayoritas (Dinas PUPR), sistem memiliki kecenderungan tinggi untuk melakukan *False Positive* atau salah disposisi. Jika model ini diimplementasikan, dampaknya adalah membanjirnya kotak masuk (inbox) admin Dinas PUPR dengan laporan-laporan yang sebenarnya bukan kewenangan mereka (misalnya laporan kependudukan atau pendidikan). Hal ini justru kontraproduktif dengan tujuan awal otomatisasi, yakni efisiensi. Alih-alih mempercepat, admin Dinas PUPR justru akan terbebani tugas tambahan untuk memilah dan mengembalikan laporan salah sasaran tersebut, yang memakan waktu lebih lama dibandingkan proses manual sebelumnya.

#### 4. Kesimpulan

Penelitian ini menghasilkan tiga kesimpulan utama: Kinerja Rendah: Model Naive Bayes hanya mencapai akurasi 24,14%, mengindikasikan kegagalan sistem dalam melakukan klasifikasi otomatis. Akar Masalah: Kegagalan ini disebabkan oleh data sparsity (jumlah sampel sangat minim, ~145 data) dan ketimpangan kelas yang ekstrem, bukan karena kelemahan algoritma.

Status Data: Dataset E-Lapor dengan volume saat ini terbukti belum layak untuk pelatihan model *machine learning* tanpa augmentasi data. Limitasi Penelitian Penelitian ini memiliki keterbatasan berupa: (1) Absennya tahapan *stemming* pada pra-pemrosesan; (2) Rentang pengambilan data yang sempit (hanya tahun 2024); dan (3) Belum adanya komparasi empiris dengan algoritma lain.

Saran Untuk perbaikan di masa mendatang, disarankan agar: Ekspansi Data: Memperluas rentang data minimal 2–3 tahun untuk mengatasi kelangkaan sampel. Optimasi Pre-processing: Wajib menerapkan *stemming* dan teknik penyeimbang data (*oversampling/SMOTE*). Komparasi

Model: Membandingkan kinerja Naive Bayes dengan algoritma lain seperti *Support Vector Machine* (SVM).

### Ucapan Terima Kasih (Opsional)

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada pihak Dinas Komunikasi, Informatika, Persandian, dan Statistik (Komdigi) Kabupaten Sumedang yang telah memberikan izin dan menyediakan data E-Lapor untuk digunakan dalam penelitian ini. Ucapan terima kasih juga penulis sampaikan kepada Bapak Agun Guntara, S.Kom., M.Kom., selaku dosen pembimbing mata kuliah, atas segala arahan, bimbingan, dan masukan berharga selama proses penyusunan penelitian dan naskah jurnal ini..

### Referensi

- [1] R. E. Berliana and I. P. D. Yudartha, "Implementasi SP4N LAPOR! dalam Meningkatkan Kualitas Pelayanan Publik (Studi Kasus: Beberapa Instansi Pemerintah yang Ada di Indonesia)," *Socio-political Commun. Policy Rev.*, vol. 1, no. 5, pp. 1–8, 2024, doi: <https://doi.org/10.61292/shkr.183>.
- [2] Hilas Stepanus and Wahyu Subadi, "Efektivitas Sistem Pengelolaan Pengaduan Pelayanan Publik Nasional - Layanan Aspirasi dan Pengaduan Online Rakyat (SP4N-LAPOR) di Kabupaten Tabalong dalam Penanganan Pengaduan Masyarakat Terkait Pelayanan Publik," *Japb*, vol. 7, no. 2, pp. 1156–1167, 2024, doi: <https://doi.org/10.35722/japb.v7i2.1070>.
- [3] I. F. Rozi, V. N. Wijayaningrum, and N. Khozin, "Klasifikasi Teks Laporan Masyarakat Pada Situs Lapor! Menggunakan Recurrent Neural Network," *Sistemasi*, vol. 9, no. 3, p. 633, 2020, doi: 10.32520/stmsi.v9i3.977.
- [4] D. Afida, E. D. Udayanti, and E. Kartikadarma, "Aplikasi Text Mining untuk Klasterisasi Aduan Masyarakat Kota Semarang Menggunakan Algoritma K-means," *J. Transform.*, vol. 18, no. 2, pp. 215–224, 2021, doi: 10.26623/transformatika.v18i2.2362.
- [5] L. Rahmawati and D. B. Santoso, "Implementasi Metode Naive Bayes Untuk Klasifikasi Ulasan Aplikasi E-Commerce Tokopedia," *J. Inf. Technol. Comput. Sci.*, vol. 6, no. 1, pp. 116–124, 2023, [Online]. Available: [https://www.researchgate.net/publication/372759197\\_Implementasi\\_Metode\\_Naive\\_Bayes\\_Untuk\\_Klasifikasi\\_Ulasan\\_Aplikasi\\_E-Commerce\\_Tokopedia](https://www.researchgate.net/publication/372759197_Implementasi_Metode_Naive_Bayes_Untuk_Klasifikasi_Ulasan_Aplikasi_E-Commerce_Tokopedia).
- [6] M. A. Adham, R. Rahmatil Fiska, and M. A. Subandri, "Metode Prototipe Aplikasi Klasifikasi Pengaduan Masyarakat Menggunakan Algoritma Naive Bayes Berbasis Website," *J. Elektron. dan Tek. Inform. Terap.*, vol. 2, pp. 61–81, 2024, doi: 10.59061/jentik.v2i4.849.
- [7] E. Yumami, M. Jannah, K. O. Putra, I. Teknologi, and M. Gama, "Klasifikasi Judul Skripsi Mahasiswa Berdasarkan Konsentrasi Program Studi Menggunakan Algoritma Naive Bayes," vol. 5, no. 1, pp. 24–28, 2025, doi: <https://doi.org/10.54840/jcstech.v5i1.352>.
- [8] C. Anam and M. Indriati, "Klasifikasi Pengaduan Pelayanan Dispendukcapil Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Glasgow-II," ... *Teknol. Inf. dan Ilmu Komput. e-ISSN*, vol. 4, no. 9, pp. 3264–3271, 2020, [Online]. Available: <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/download/7932/3728>.
- [9] Y. Sulistiowati and B. J. Santoso, "Analisis Sentimen Ulasan Pengguna Aplikasi Mobile SP4N-LAPOR! dengan Pendekatan Machine Learning," *J. Inform. Polinema*, vol. 11, no. 3, pp. 283–290, 2025, doi: 10.33795/jip.v11i3.7189.
- [10] W. Clarisha, A. A. M. Fani, D. F. Suriyanto, and N. Fadilah, "Analisis Sentimen Sunscreen Lokal Skintific, Somethinc, dan Avoskin dengan Naive Bayes dan SVM," *Jambura J. Electr. Electron. Eng.*, vol. 7, no. 2, pp. 264–271, 2025, [Online]. Available: <https://ejurnal.ung.ac.id/index.php/jjeee/article/viewFile/30257/11225>.
- [11] K. Mumbunan, M. Marchantia Betsi Bawata, M. Prayer kusen, V. Tarigan, and A. Yusupa, "Perbandingan Algoritma Naive Bayes Dan Support Vector Machine Dalam Deteksi Berita Hoax Berbahasa Indonesia," *Riau J. Tek. Inform.*, vol. 4, no. 1, pp. 52–64, 2025, doi: 10.30606/rjti.v4i1.3262.
- [12] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.

- [13] S. Chamira, "Implementasi Metode Text Mining Frequency-Invers Document Frequency (Tf-Idf) Untuk Monitoring Diskusi Online," *J. Informatics, Electr. Electron. Eng.*, vol. 1, no. 3, pp. 97–102, 2022, doi: 10.47065/jieeee.v1i3.353.
- [14] M. H. Z. Ubaidillah and Z. Fatah, "Klasifikasi Email Spam Menggunakan Algoritma Naive Bayes," *Gudang J. Multidisiplin Ilmu*, vol. 2, no. November, pp. 129–132, 2024, doi: 10.59435/gjmi.v2i11.536.
- [15] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, pp. 1–21, 2021, doi: 10.1007/s42979-021-00592-x.