UNIVERSITAS
MUHAMMADIYAH
SIDOARJO

UMSIDA
DARI SINI PENCERAHAN BERSEMI

Sentiment Analysis Of Public Perceptions Of The President's Policy On The National Disaster Status In Sumatra Using A Social Media-Based Naive Bayes Model X

Oleh :

Naufal Farras Hilmya Hanif

Dosen Pembimbing:

Hamzah Setiawan, S.Kom., M.Kom.

Dosen Penguji 1:

Arif Senja Fitrani, S.Kom., M.Kom.

Dosen Penguji 2:

Ade Eviyanti, S.Kom., M.Kom.

Program Studi Informatika

Universitas Muhammadiyah Sidoarjo

2026

Pendahuluan

Latar Belakang Masalah

- Bencana banjir di kawasan Sumatera menuntut penetapan status darurat nasional oleh Presiden, yang memantik respons luas dari masyarakat.
- Media Sosial X (Twitter) menjadi ruang utama bagi publik untuk menyuarakan opini secara terbuka dan *real-time*.
- Dibutuhkan pendekatan komputasional (*Machine Learning*) untuk memetakan volume opini publik yang masif secara akurat.
- Penelitian terdahulu mengenai sentimen publik terhadap intervensi krisis di tingkat negara masih sangat minim.

Pendahuluan

Rumusan Masalah

- Bagaimana sentimen dan persepsi publik di platform X terhadap kebijakan Presiden terkait status bencana nasional banjir di Sumatera?
- Bagaimana mengatasi kendala *imbalanced data* agar model dapat mendeteksi kelas sentimen minoritas dengan akurat?

Pendahuluan

Tujuan Penelitian

- Menganalisis sentimen masyarakat terhadap kebijakan Presiden mengenai status bencana nasional banjir di Sumatera menggunakan metode Multinomial Naïve Bayes.
- Menerapkan pembobotan TF-IDF dan teknik *oversampling* untuk mendapatkan performa model yang optimal.

Pendahuluan

Batasan Masalah

- Data opini diambil dari platform Media Sosial X dengan 5 kata kunci spesifik.
- Pelabelan sentimen otomatis dilakukan menggunakan pendekatan leksikon VADER.
- Algoritma dibatasi pada Multinomial Naïve Bayes dengan fitur TF-IDF.
- *Oversampling* diaplikasikan eksklusif pada *training data* untuk menyeimbangkan kelas.

Metodologi

Alur Tahapan

Penelitian menggunakan pendekatan kuantitatif komputasional dalam ranah *Natural Language Processing* (NLP) dengan tahapan utama:

1. Studi Literatur
2. *Scraping* Data
3. Pelabelan (*Labeling*)
4. *Oversampling*
5. *Pre-Processing*
6. Pembagian Data (*Split Data*)
7. Pembobotan TF-IDF
8. Perbandingan Naïve Bayes
9. Evaluasi Model

Metodologi

Pengumpulan & Pelabelan Data

- **Scraping Data:** Dilakukan menggunakan teknik *web scraping* dengan pustaka Selenium pada Python, yang menghasilkan 1.321 data teks cuitan mentah.
- **Pelabelan Data:** Menggunakan leksikon VADER yang mencocokkan setiap kata dengan kamus untuk menghasilkan nilai komposit (*compound score*).
- **Distribusi Awal:** Dari 1.208 cuitan bersih yang valid, kelas Netral sangat mendominasi (92,6%), sedangkan kelas Positif (5,0%) dan Negatif (2,4%) sangat minim.

Metodologi

Pre-Processing & Penanganan Ketidakseimbangan Data

- **Pembersihan Teks (Pre-Processing):** Meliputi *Case Folding* (huruf kecil), *Cleansing* (hapus URL/emoji), Normalisasi (kata baku), *Tokenizing*, *Stopword Removal*, dan *Stemming* menggunakan Sastrawi.
- **Oversampling:** Untuk memitigasi risiko bias komputasi terhadap kelas mayoritas, data pada kelas minoritas disintesis secara proporsional khusus pada *training data*.
- **Hasil Oversampling:** Total data berhasil ditingkatkan menjadi 2.685 data dengan distribusi seimbang sempurna khusus data latih saja (masing-masing 2.685 data untuk sentimen Netral, Positif, dan Negatif).

Metodologi

Ekstraksi Fitur & Evaluasi Model

- **Split Data:** Teks yang telah bersih dipartisi secara acak menjadi 80% data latih dan 20% data uji.
- **Pembobotan TF-IDF:** Mengubah representasi teks menjadi vektor numerik untuk mengukur signifikansi lokal sebuah kata sekaligus memberi penalti pada kata yang terlalu sering muncul.
- **Klasifikasi Naïve Bayes:** Vektor hasil ekstraksi diumpankan ke dalam tiga algoritma untuk perbandingan yaitu Multinomial Naïve Bayes, Bernoulli Naïve Bayes, Complement Naïve Bayes
- **Evaluasi:** Kinerja prediksi terhadap data uji dinilai menggunakan instrumen *Confusion Matrix* yang menghasilkan empat metrik pengukuran, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score*.

Hasil dan Pembahasan



Dataset

Berjumlah 2.685 data tweet pada Twitter/*X* setelah validasi pembersihan tweet yang valid dan proses *Oversampling*, rentang waktu pengambilan:

14 Desember 2025 – 26 Mei 2026

Distribusi Sentimen

Total 1.208 data dilabeli dengan leksikon VADER (60 Positif, 29 Negatif, 1.119 Netral). Setelah melalui 6 tahap *preprocessing*, diterapkan teknik *oversampling* untuk menyeimbangkan data latih menjadi 2.685 data (895 data per kelas sentimen). Selanjutnya dilakukan *split data* dengan rasio 80:20, pembobotan teks ke numerik melalui TF-IDF, serta pelatihan dan perbandingan tiga algoritma: Multinomial, Bernoulli, dan Complement Naïve Bayes. Evaluasi performa ketiga model tersebut kemudian diukur menggunakan *Confusion Matrix*, *Accuracy*, *Precision*, *Recall*, dan *F1-Score*

Split Data

- **Pembagian Data:** Dataset dipartisi menjadi data latih dan data uji menggunakan metode *stratified random split* dengan rasio pembagian 80:20.
- **Penanganan *Imbalanced Data*:** Untuk mengatasi ketimpangan kelas ekstrem (sentimen Netral mendominasi 92,6%), diterapkan teknik penyeimbangan data menggunakan metode *Oversampling*.
- **Proses *Oversampling*:** Sesuai kaidah *Machine Learning*, teknik ini diaplikasikan eksklusif pada data latih dengan mereplikasi kelas minoritas (Positif dan Negatif) hingga setara dengan kelas mayoritas, yakni masing-masing menjadi 895 data.
- **Alokasi Pelatihan & Pengujian:** Data latih (80% atau total 2.685 data) digunakan untuk melatih tiga model (*Multinomial*, *Bernoulli*, dan *Complement Naïve Bayes*), sedangkan data uji (20% atau 242 data) dikhususkan murni untuk mengevaluasi kinerja prediksi model.
- **Tujuan:** Memastikan algoritma mempelajari karakteristik setiap kelas secara adil tanpa bias kelas mayoritas selama pelatihan, sekaligus menjaga agar pengujian tetap menggunakan kondisi data riil yang objektif

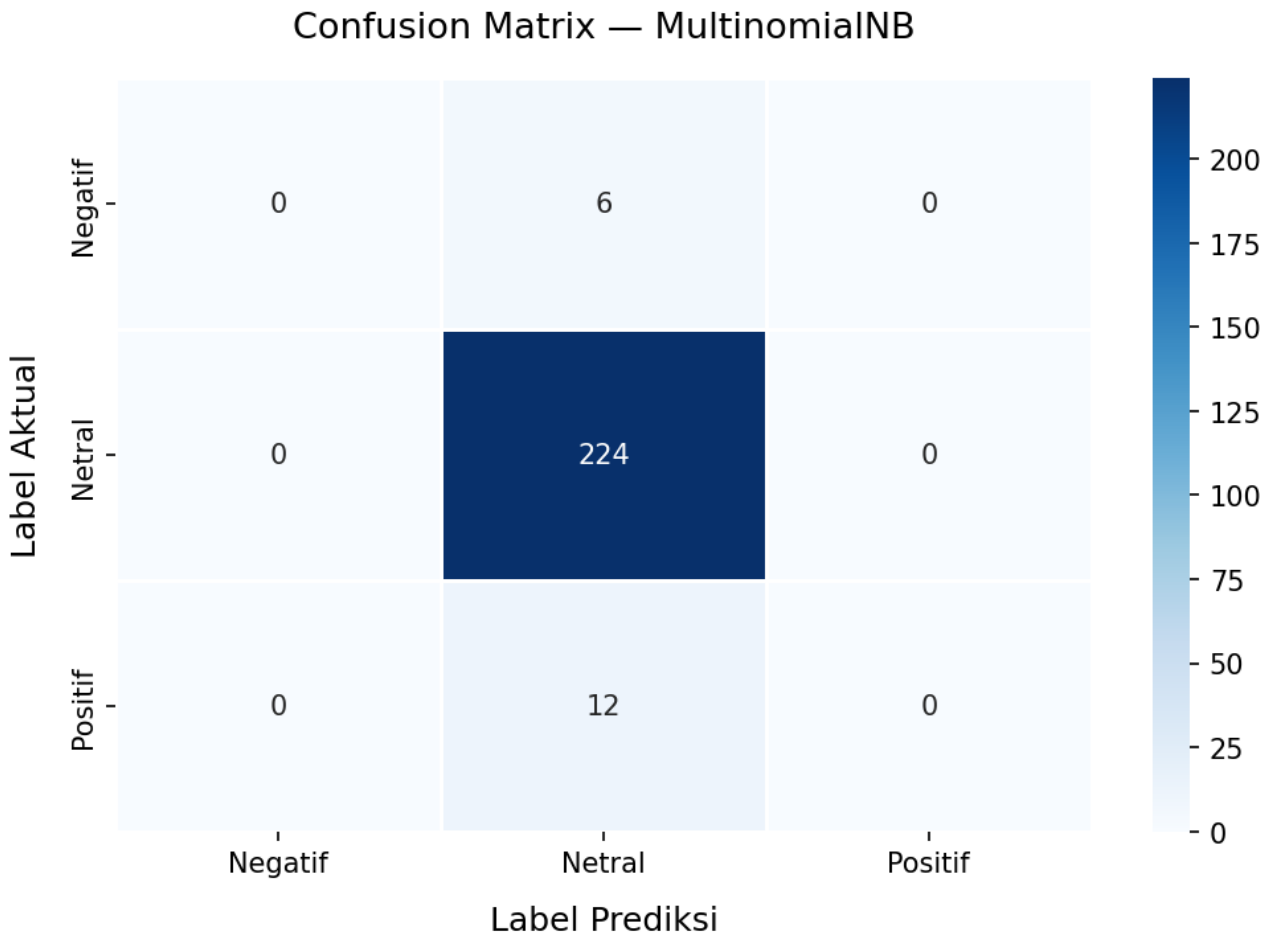
Split Data

Tabel Distribusi Pembagian Data (Rasio 80:20)

Kategori Sentimen	Data Latih (Training) (Setelah Oversampling)	Data Uji (Testing) (Data Asli)
Netral	895	224
Positif	895	12
Negatif	895	6
Total	2.685 (80%)	242 (20%)

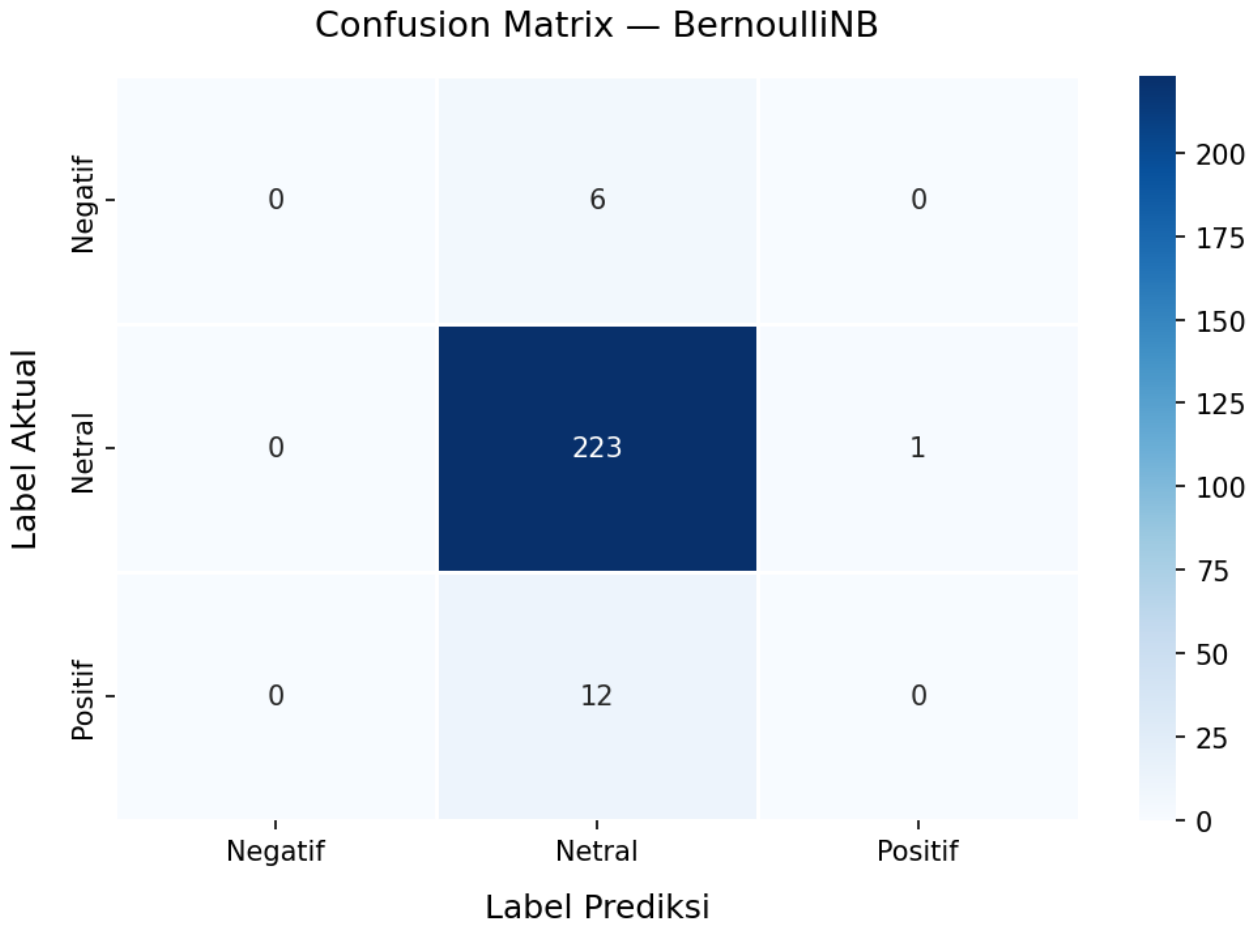
- Dataset awal dipartisi menggunakan metode *Stratified Random Split* dengan rasio pembagian 80:20.
- Sebanyak 80% data dialokasikan sebagai data latih, yang kemudian dilakukan *oversampling* menjadi total 2.685 data seimbang guna melatih tiga model klasifikasi: **Multinomial, Bernoulli, dan Complement Naïve Bayes**.
- Sebanyak 20% (242 data) murni disisihkan sebagai data uji untuk mengevaluasi kinerja prediksi model pada data baru. Sesuai kaidah *Machine Learning*, data uji ini **tidak di-oversampling** agar evaluasi tetap objektif pada kondisi riil.
- Penggunaan metode *Stratified* memastikan setiap kelas sentimen (Positif, Negatif, Netral) terbagi secara proporsional ke dalam set data latih maupun data uji.

Naïve Bayes 80% Latih : 20 % Uji



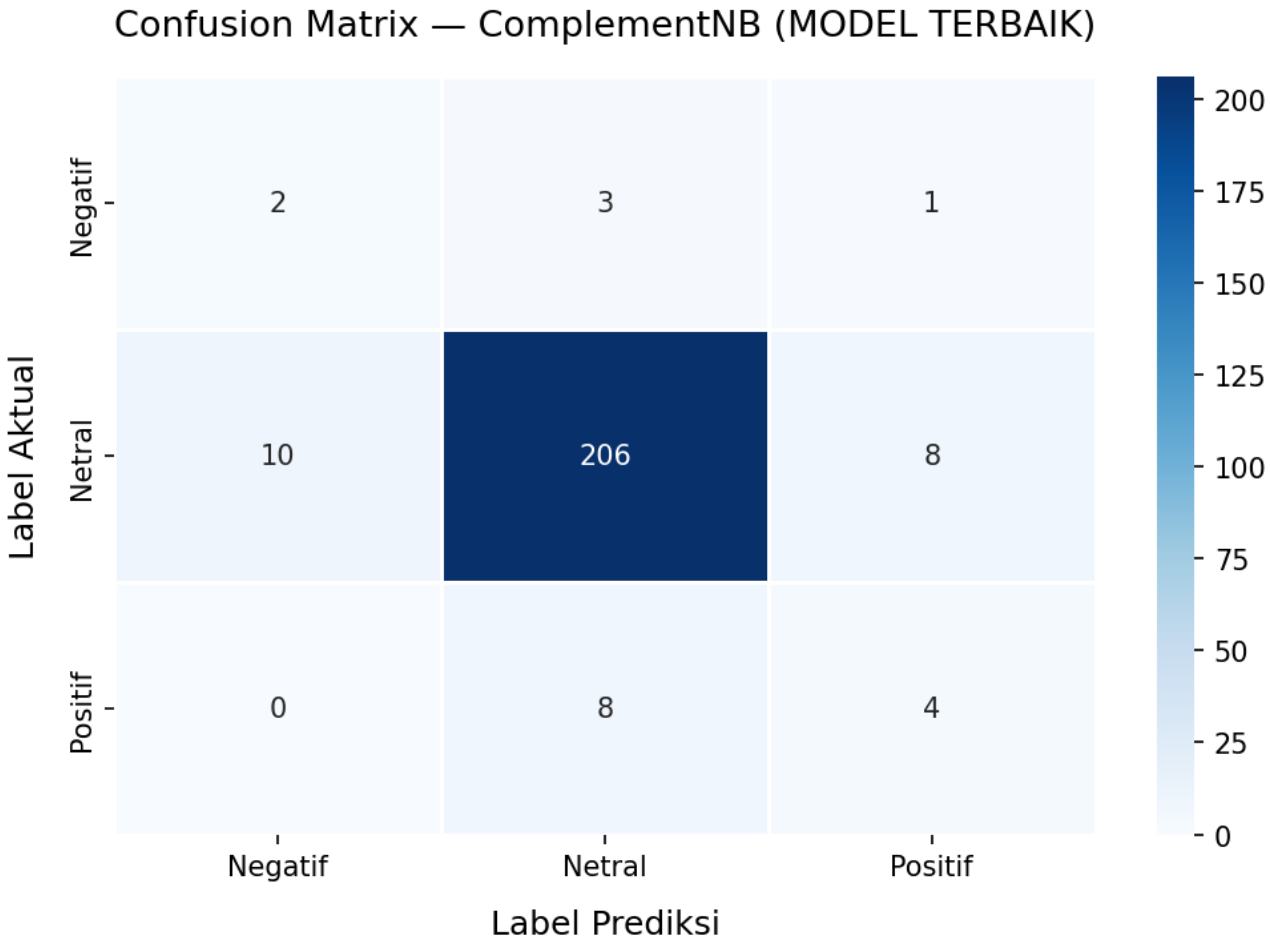
Pada skenario pembagian data 80:20, model Multinomial Naïve Bayes mencatatkan tingkat akurasi (*Accuracy*) keseluruhan sebesar 93%. Namun, tingginya angka akurasi ini merupakan sebuah paradoks karena model terbukti gagal total dalam mendeteksi sentimen pada kelas minoritas. Berdasarkan hasil pada *confusion matrix*, model ini hanya berhasil mengklasifikasikan sentimen kelas mayoritas secara tepat, yakni 224 data Netral. Sebaliknya, model ini mengalami *overfitting* yang parah pada kelas Netral, terbukti dari seluruh data kelas minoritas yang mengalami misklasifikasi: 6 data yang sebenarnya Negatif keliru diprediksi menjadi Netral, dan 12 data yang sebenarnya Positif juga keliru diprediksi menjadi Netral.

Naïve Bayes 80% Latih : 20 % Uji



Pada skenario pembagian data 80:20, model Bernoulli Naïve Bayes mencatatkan tingkat akurasi (*Accuracy*) keseluruhan sebesar 92%. Namun, serupa dengan model sebelumnya, tingginya angka akurasi ini merupakan sebuah paradoks karena model terbukti gagal total dalam mendeteksi sentimen pada kelas minoritas (positif dan negatif). Berdasarkan hasil pada *confusion matrix*, model ini hampir seluruhnya hanya memprediksi kelas mayoritas secara tepat, yakni sebanyak 223 data Netral. Model ini mengalami *overfitting* yang parah pada sentimen Netral, terbukti dari seluruh data kelas minoritas yang mengalami misklasifikasi: 6 data yang sebenarnya Negatif keliru diprediksi menjadi Netral, dan 12 data yang sebenarnya Positif juga keliru diprediksi menjadi Netral. Selain itu, terdapat 1 data Netral yang keliru diprediksi menjadi kelas Positif.

Naïve Bayes 80% Latih : 20 % Uji



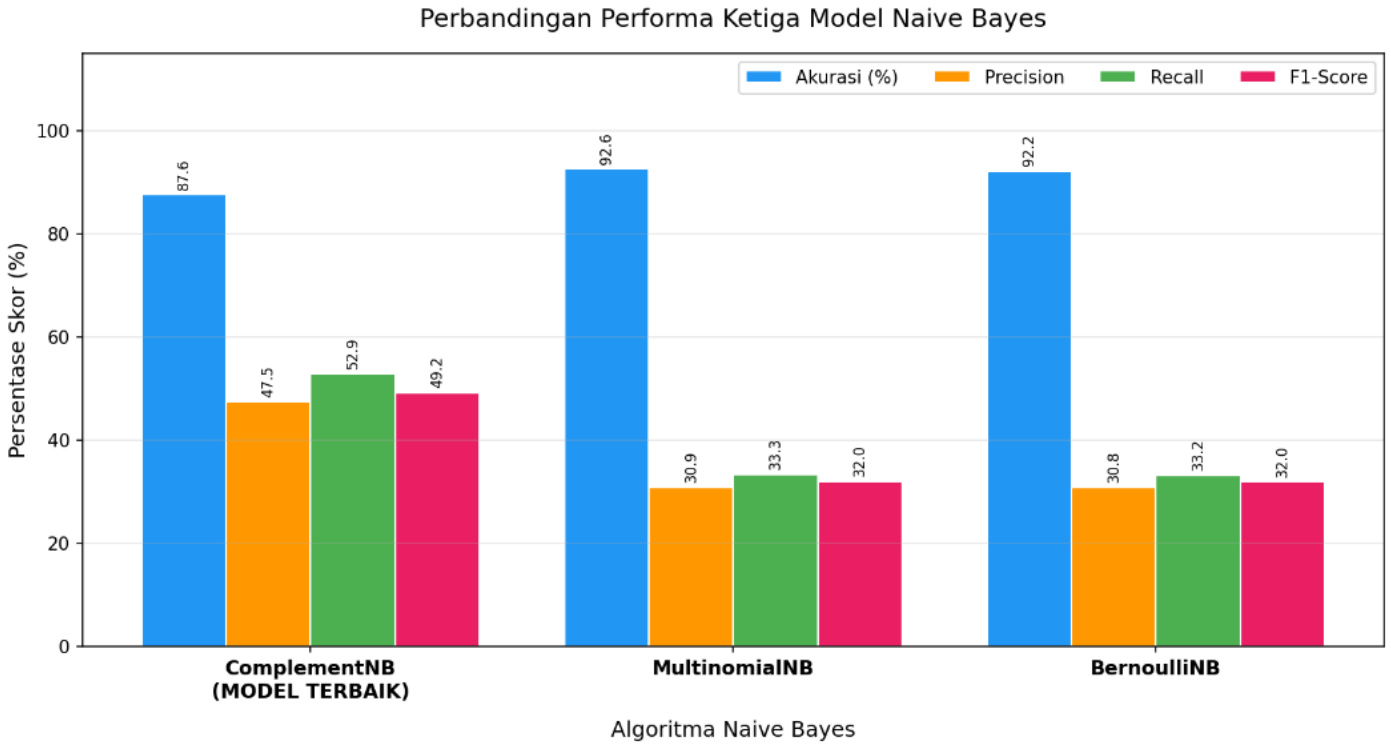
Pada skenario pembagian data 80:20, model Complement Naïve Bayes mencatatkan tingkat akurasi (*Accuracy*) keseluruhan sebesar 87,60% (dibulatkan menjadi 88%). Berbeda dengan dua algoritma sebelumnya yang mengalami *overfitting* ekstrem, ComplementNB terbukti menjadi **Model Terbaik** yang paling andal dan adaptif dalam menangani data yang sangat tidak seimbang.

Berdasarkan hasil pada *confusion matrix*, ketangguhan ini dibuktikan dengan keberhasilan model dalam mendeteksi sentimen kelas minoritas. Model ini berhasil memprediksi dengan tepat 206 data Netral, 4 data Positif, dan 2 data Negatif. Meskipun masih terdapat sebagian misklasifikasi seperti 8 data Positif dan 3 data Negatif yang keliru diprediksi sebagai Netral, serta beberapa data Netral yang terprediksi sebagai kelas minoritas (10 menjadi Negatif, 8 menjadi Positif) model ini membuktikan bahwa ia tidak sepenuhnya terdistorsi oleh bias kelas mayoritas, sehingga menjadi solusi klasifikasi yang paling optimal dan proporsional

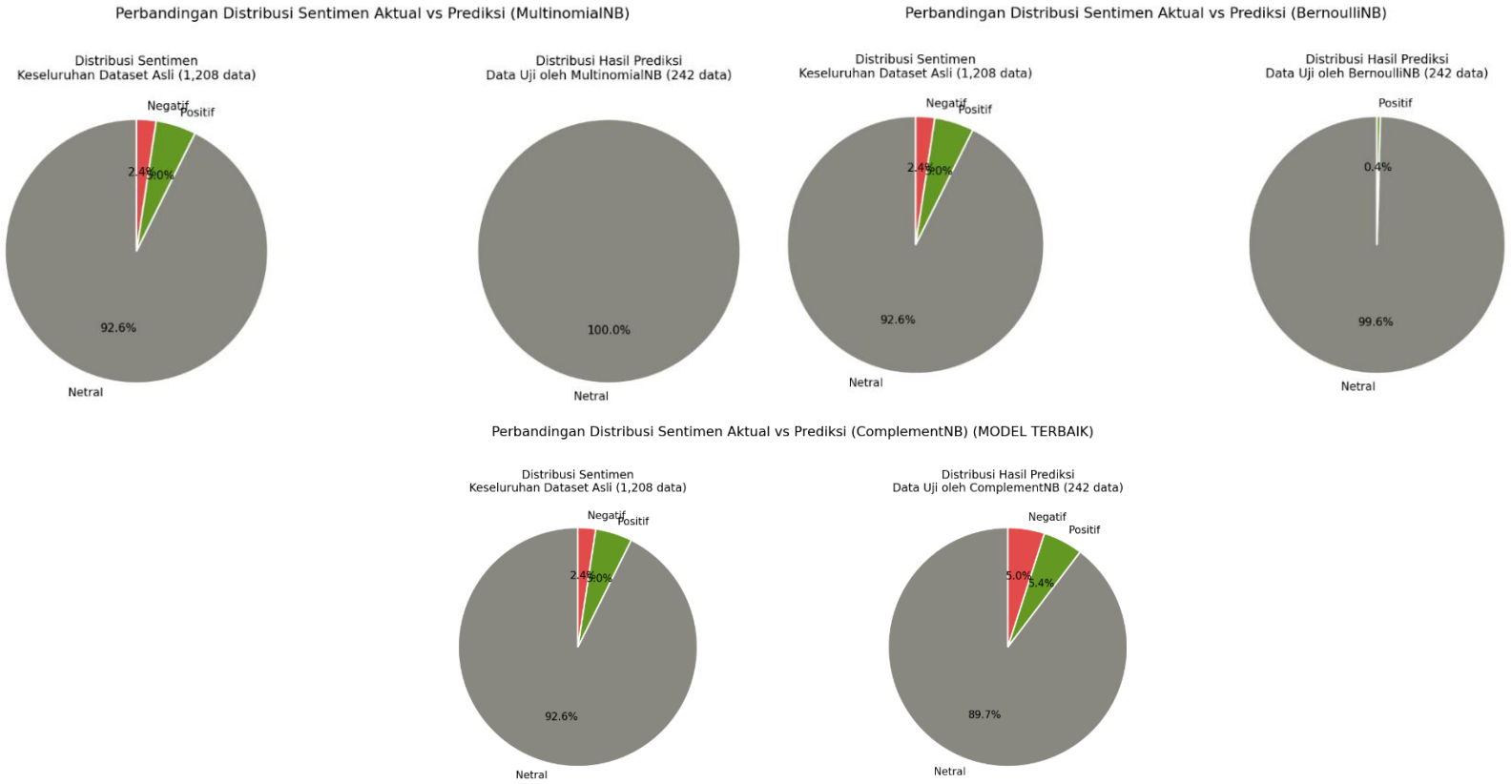
Evaluasi Perbandingan Model Per Kelas Sentimen

Perbandingan Performa Ketiga Model Naïve Bayes

- **Paradoks Akurasi (MultinomialNB & BernoulliNB):** Meskipun MultinomialNB (92,6%) dan BernoulliNB (92,2%) memiliki akurasi yang sangat tinggi, hal ini merupakan sebuah paradoks. Kedua model tersebut mengalami *overfitting* ekstrem pada kelas mayoritas (Netral) dan gagal total mendeteksi kelas minoritas. Akibatnya, nilai rata-rata *Precision*, *Recall*, dan *F1-Score* mereka sangat rendah, yakni berada di kisaran 30%-33%.
- **Model Terbaik (ComplementNB):** Model Complement Naïve Bayes terbukti sebagai algoritma yang paling andal dan adaptif dalam menangani kasus ketimpangan data ini, dengan tingkat akurasi yang solid sebesar 87,6%.
- **Keseimbangan Performa:** ComplementNB memiliki sensitivitas dan ketepatan paling tangguh untuk mengungkap polaritas sentimen minoritas tanpa terdistorsi oleh dominasi kelas mayoritas. Keberhasilan ini ditegaskan oleh perolehan nilai tertinggi pada *Precision* (47,5%), *Recall* (52,9%), dan *F1-Score* (49,2%).



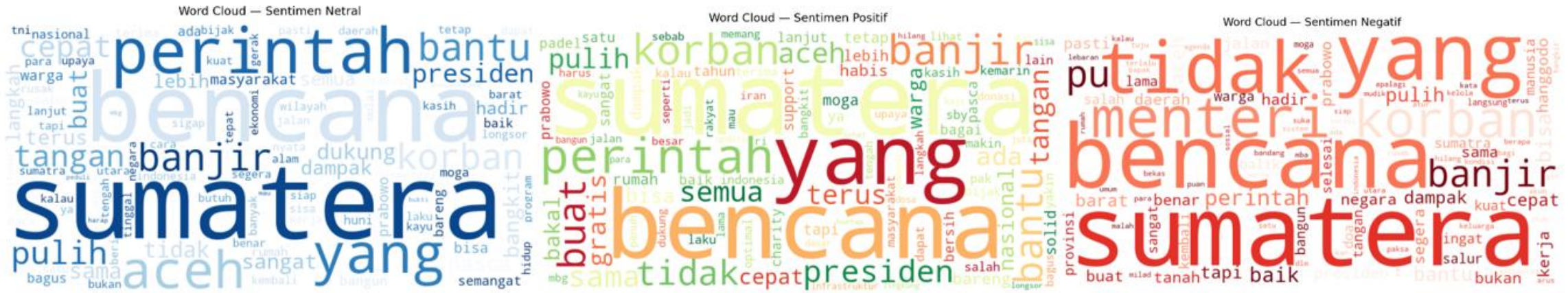
Perbandingan Distribusi Sentimen



Perbandingan Distribusi Sentimen Aktual vs Prediksi

- **Distribusi Awal (Aktual):** Distribusi sentimen pada 1.208 data keseluruhan didominasi sangat kuat oleh kelas Netral sebesar 92,6%, berbanding terbalik dengan kelas Positif (5,0%) dan Negatif (2,4%).
- **Multinomial Naïve Bayes:** Model memprediksi 100% keseluruhan data uji masuk ke dalam sentimen Netral. Hal ini mengindikasikan model mengalami *overfitting* yang parah dan sama sekali gagal mengungkap polaritas sentimen minoritas.
- **Bernoulli Naïve Bayes:** Daya prediksi model tetap terjebak pada bias kelas mayoritas dengan memprediksi 99,6% sentimen Netral, hanya menyisakan porsi yang sangat marjinal untuk Positif (0,4%) dan nihil pada sentimen Negatif.
- **Complement Naïve Bayes (Model Terbaik):** Terbukti secara algoritmik lebih tangguh dan adaptif. Model ini mampu menyusutkan proporsi dominasi Netral menjadi 89,7%, sehingga berhasil mengungkap polaritas sentimen minoritas yang sebelumnya tertekan, yakni sentimen Positif sebesar 5,4% dan Negatif sebesar 5,0%.

Analisis Word Cloud



- **Sentimen Netral**: Didominasi oleh kata-kata ekstensif seperti "sumatera", "banjir", "bencana", "aceh", "perintah", "bantu", dan "presiden". Cuitan pada kelas ini bersifat informatif, pelaporan, dan lugas yang berfokus pada pembaruan situasi terkini pascabencana, penyaluran bantuan, serta pelaporan objektif mengenai kesigapan aparat di lapangan tanpa muatan emosional berlebihan.
- **Sentimen Positif**: Didominasi oleh kata-kata kunci sentral seperti "sumatera", "bencana", "banjir", "bantu", "perintah", "presiden", dan "aceh" yang disertai kata-kata bernada optimis dan suportif seperti "bangkit", "pulih", "cepat", "bagus", "solid", serta "optimal". Pola ini mengindikasikan apresiasi, dukungan moril, solidaritas, dan harapan agar proses pemulihan serta penanganan korban berjalan cepat dan terkoordinasi.
- **Sentimen Negatif**: Didominasi oleh kata-kata menonjol seperti "sumatera", "bencana", "menteri", "pu", "perintah", dan "korban". Terdapat pula kata-kata spesifik yang mengarah pada keluhan dan kritik seperti "tidak", "salah", "lama", "rusak", "dampak", serta "jalan", termasuk kemunculan nama spesifik yang menegaskan kritik tajam terhadap kinerja Kementerian PU dan kelambanan penanganan pascabencana.

Kesimpulan

Penelitian ini berhasil menggunakan 3.357 data teks seimbang hasil *oversampling* (terdiri dari masing-masing 1.119 sentimen positif, negatif, dan netral dari total 1.208 cuitan valid awal), serta menguji algoritma Multinomial *Naïve Bayes* dengan pembobotan TF-IDF pada rasio pembagian data 80:20. Hasil pengujian menunjukkan bahwa model ini berkinerja sangat andal dengan *Accuracy* mencapai 94,79%, *Precision* 0,9492, *Recall* 0,9479, dan *F1-Score* 0,9479. Keunggulan utama dari model yang dikombinasikan dengan teknik *oversampling* ini terletak pada kemampuannya mendeteksi polaritas opini publik secara proporsional (memprediksi 35,6% positif, 33,3% negatif, dan 31,1% netral pada data uji) tanpa terdistorsi oleh bias kelas mayoritas. Sementara itu, analisis *WordCloud* menunjukkan bahwa sentimen negatif didominasi oleh kritik terhadap kelambanan birokrasi kementerian dan kerusakan infrastruktur, sentimen netral berisi sirkulasi informasi faktual terkait program rehabilitasi dan laporan aparat, dan sentimen positif mencerminkan apresiasi publik terhadap langkah taktis pemulihan serta distribusi bantuan pemerintah. Untuk penelitian selanjutnya disarankan memperluas pengambilan data ke platform media sosial lain, menambahkan leksikon bahasa daerah atau bahasa gaul lokal, serta mengeksplorasi perbandingan dengan model algoritma lain seperti SVM, *Random Forest*, maupun algoritma *deep learning* seperti LSTM.


UMSIDA
DARI SINI PENCERAHAN BERSEMI