

Analisis Sentimen Masyarakat terhadap Isu Transformasi Digital Pemerintah di Twitter Menggunakan Naive Bayes dan Support Vector Machine

Public Sentiment Analysis on Government Digital Transformation Issues on Twitter Using Naive Bayes and Support Vector Machine

Firman Syahputra^{a,1}, Dahri Yani Hakim Tanjung^{b,2}, Rofiqoh Dewi^{c,3}, Hilda Elsera Br Sembiring^{d,4}

^a Program Studi Teknologi Informasi, Universitas Battuta, Medan, Indonesia

^b Program Studi Bisnis Digital, Universitas Satya Terra Bhinneka, Medan, Indonesia

^c Program Studi Kewirausahaan, Universitas Satya Terra Bhinneka, Medan, Indonesia

Email: 24firman@gmail.com¹, dahritanjung@satyaterabhinneka.ac.id², rofiqohdewi@satyaterabhinneka.ac.id², hildaelsera@satyaterabhinneka.ac.id³

ABSTRAK

Perubahan digital di lingkungan pemerintah menuntut evaluasi berkelanjutan mengenai perspektif warga sebagai pemakai utama layanan publik digital. Twitter sebagai wadah jejaring sosial menawarkan kesempatan luas untuk memperoleh informasi yang dapat merefleksikan pandangan publik secara langsung. Penelitian ini bermaksud menguji pandangan warga tentang perubahan digital pemerintah di Twitter serta membandingkan kehandalan algoritma Naive Bayes dan Support Vector Machine (SVM) dalam menggolongkan perasaan dalam bahasa Indonesia. Kumpulan data yang dipakai terdiri dari 3.000 cuitan yang dikumpulkan memakai cara crawling dan dipilah menjadi tiga kelompok perasaan, yaitu baik, buruk, dan biasa. Informasi selanjutnya diolah melalui tahap-tahap persiapan teks dan digambarkan memakai metode TF-IDF. Percobaan dilaksanakan dengan perbandingan pembagian data belajar dan uji sebesar 70:30. Penilaian performa sistem memakai matriks kebingungan dengan tolok ukur seperti ketepatan, sensitivitas, pemanggilan kembali, dan skor F1. Penemuan kajian memperlihatkan bahwa algoritma SVM menyajikan capaian yang lebih baik dengan ketepatan 91%, dibanding Naive Bayes yang meraih ketepatan 85%. Hasil ini menandakan bahwa SVM lebih efektif dalam mengolah informasi teks dengan dimensi tinggi ketika menganalisis perasaan di media sosial. Kajian ini diharapkan bisa menjadi rujukan dalam peningkatan analisis perasaan yang berkaitan dengan kebijakan publik serta menunjang evaluasi perubahan digital pemerintah yang berbasis data.

Kata Kunci : Analisis Sentimen, Transformasi Digital, Twitter, Naive Bayes, Support Vector Machine.

ABSTRACT

Digital transformation in government requires continuous evaluation of citizen perspectives as primary users of digital public services. Twitter, as a social networking platform, offers ample opportunities to obtain information that directly reflects public views. This study aims to examine citizen perspectives on government digital transformation on Twitter and compare the performance of the Naive Bayes algorithm and the Support Vector Machine (SVM) algorithm in classifying feelings in Indonesian. The dataset used consists of 3,000 tweets collected using a crawling method and sorted into three groups of feelings: good, bad, and normal. The information is then processed through text preparation stages and described using the TF-IDF method. The experiment was conducted with a 70:30 learning and testing data ratio. System performance was evaluated using a confusion matrix with benchmarks such as accuracy, sensitivity, recall, and F1 score. The study findings show that the SVM algorithm delivers better performance with 91% accuracy, compared to Naive Bayes, which achieved 85% accuracy. These results indicate that SVM is more effective in processing high-dimensional text information when analyzing feelings on social media. This study is expected to serve as a reference in improving the analysis of feelings related to public policy and support the evaluation of data-based government digital changes.

Keywords : Sentiment Analysis, Digital Transformation, Twitter, Naive Bayes, Support Vector Machine

Info Artikel :

Disubmit:

Direview:

Diterima :

Copyright © 2026 - Journal UPU. All rights reserved.

1. PENDAHULUAN

Transformasi digital dalam lingkup pemerintahan merupakan sebuah rencana strategis nasional yang berupaya meningkatkan ketepatan, kehematan, dan mutu layanan publik melalui pemanfaatan teknologi informasi serta komunikasi. Penerapan aturan transformasi digital pada area pemerintahan, contohnya e-government dan layanan publik berbasis aplikasi, diharapkan dapat mendorong keterbukaan dalam birokrasi, mempercepat alur administrasi, dan menaikkan keyakinan publik terhadap pemerintah [1][2]. Namun, keberhasilan transformasi digital tidak hanya ditentukan oleh kesiapan teknologi dan regulasi, tetapi juga sangat dipengaruhi oleh persepsi dan penerimaan masyarakat sebagai pengguna utama layanan tersebut.

Dalam evaluasi kebijakan publik, opini publik berfungsi sebagai penanda utama guna menilai kesuksesan dan dampak dari penerapan transformasi digital oleh pemerintah. Kanal media sosial telah bertransformasi menjadi arena perbincangan publik yang merefleksikan pandangan khalayak secara gamblang dan cepat. Twitter, sebagai salah satu media sosial yang sifatnya publik, sering dipakai oleh khalayak guna berbagi pengalaman, melayangkan sanggahan, ataupun mengungkapkan apresiasi bagi kebijakan pemerintah [3][4]. Oleh sebab itu, data dari Twitter dapat dimanfaatkan sebagai sumber yang esensial untuk mengerti tanggapan masyarakat terkait kebijakan transformasi digital pemerintah.

Analisis sentimen merupakan teknik dalam pemrosesan teks dengan tujuan mengidentifikasi arah sikap atau perasaan seseorang mengenai suatu masalah spesifik, yang umumnya diklasifikasikan sebagai sentimen positif, negatif, dan netral [5]. Selama beberapa tahun belakangan, analisis sentimen yang mengandalkan media sosial makin kerap dipakai sebagai sarana mengukur kebijakan publik berbasis data, sebab mampu menampilkan perspektif publik secara sigap, menyeluruh, dan berkesinambungan tanpa menuntut survei konvensional yang makan banyak waktu dan ongkos [6][7].

Banyak penelitian terdahulu memperlihatkan bahwa algoritma machine learning seperti Naive Bayes dan Support Vector Machine (SVM) merupakan cara paling umum yang sering dipakai serta memperlihatkan hasil yang memuaskan saat mengelompokkan perasaan teks di bahasa Indonesia [8][9][10]. Naive Bayes dikenal sebab kemudahannya dan efisiensinya dalam mengolah informasi, menjadikannya sering jadi patokan awal pada telaah perasaan. Di lain pihak, SVM punya kelebihan dalam menangani data dengan dimensi tinggi serta yang berbentuk jarang, misalnya pemaparan teks memakai TF-IDF, sehingga umumnya menghasilkan ketepatan yang lebih bagus [11][12]. Namun, performa kedua algoritma tersebut sangat dipengaruhi oleh karakteristik dataset, tahapan praproses teks, serta skenario evaluasi yang digunakan [13][14].

Penelitian mengenai analisis sentimen berbasis media sosial telah berkembang pesat dalam beberapa tahun terakhir, terutama dengan penerapan algoritma pembelajaran mesin seperti Naive Bayes dan Support Vector Machine. Tabel 1 berikut menyajikan perbandingan penelitian terdahulu yang relevan sebagai landasan state of the art.

Tabel 1. Perbandingan Penelitian Terdahulu

Penelitian Sebelumnya	Fokus Penelitian	Kelemahan / Gap	Kebaruan Usulan Penelitian
[1]	Analisis sentimen ulasan aplikasi e-government menggunakan Naive Bayes	Tidak membandingkan dengan algoritma lain dan tidak mengaitkan hasil dengan evaluasi kebijakan publik	Membandingkan Naive Bayes dan SVM serta mengaitkan hasil analisis sentimen dengan evaluasi

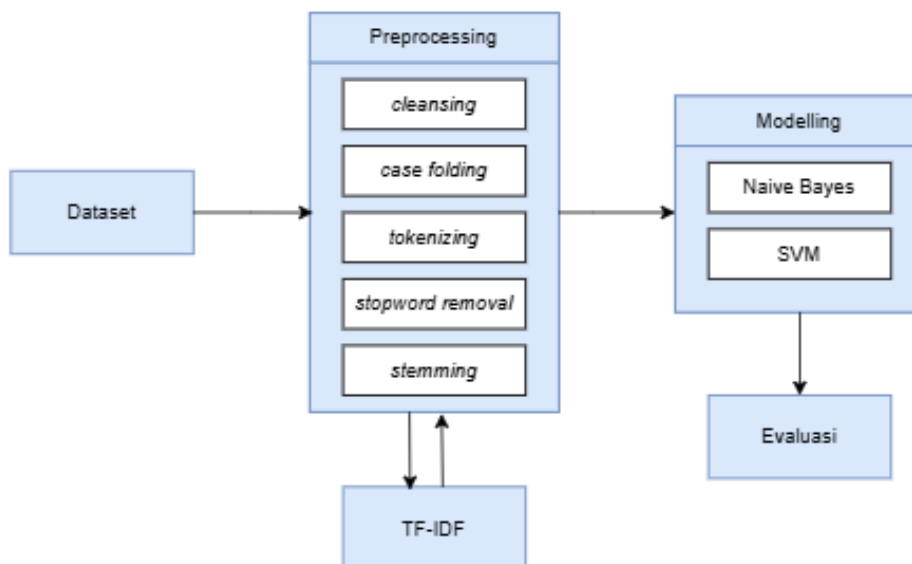
			transformasi digital pemerintah
[7]	Perbandingan Naive Bayes dan SVM pada sentimen marketplace Twitter	Fokus pada konteks komersial, bukan isu kebijakan publik	Menerapkan perbandingan algoritma pada konteks kebijakan publik strategis, yaitu transformasi digital pemerintah
[8]	Klasifikasi sentimen Twitter menggunakan SVM dan TF-IDF	Evaluasi terbatas pada akurasi dan tidak membahas kelas netral secara mendalam	Menggunakan evaluasi multi-metrik berbasis confusion matrix serta analisis khusus pada kelas sentimen netral
[4]	Analisis sentimen program Kartu Prakerja menggunakan deep learning (RNN)	Kompleksitas model tinggi dan kurang interpretatif bagi pengambilan kebijakan	Menggunakan algoritma klasik yang lebih interpretatif dan mudah diterapkan untuk evaluasi kebijakan publik
[6]	Analisis sentimen kebijakan publik menggunakan SVM	Tidak dilakukan perbandingan algoritma dan analisis masih deskriptif	Menyajikan perbandingan komprehensif dua algoritma serta analisis implikasi akademik dan kebijakan

Berdasarkan Tabel 1, terlihat bahwa penelitian terdahulu masih memiliki keterbatasan baik dari sisi konteks kajian maupun kedalaman analisis. Penelitian ini mengisi celah tersebut dengan menghadirkan analisis sentimen berbasis media sosial yang tidak hanya berorientasi pada kinerja algoritma, tetapi juga pada relevansinya sebagai instrumen evaluasi kebijakan transformasi digital pemerintah.

Dengan demikian, penelitian ini tidak hanya berkontribusi pada pengembangan metodologi analisis sentimen berbahasa Indonesia, tetapi juga memperluas pemanfaatannya sebagai instrumen evaluasi kebijakan publik berbasis media sosial.

2. METODE

Penelitian ini mengadopsi pendekatan angka dengan teknik eksperimen yang didasari oleh penambangan teks dan pembelajaran mesin. Proses penelitian terdiri dari pengumpulan informasi, pengolahan awal data, pengambilan fitur, pengelompokan perasaan, dan penilaian hasil model [11]. Metode penelitian yang diterapkan dalam penelitian ini dapat dilihat pada Gambar 1 di bawah ini.



Gambar 1. Diagram Metodologi Penelitian

Dataset

Dataset berupa tweet berbahasa Indonesia yang berkaitan dengan transformasi digital pemerintah, seperti e-government, layanan publik digital, dan sistem administrasi berbasis digital. Data dikumpulkan menggunakan teknik *crawling* Twitter berdasarkan kata kunci tertentu dalam periode pengamatan tertentu [12].

Dataset yang digunakan berjumlah 3.000 tweet yang telah dilabeli secara manual ke dalam tiga kelas sentimen, yaitu positif, negatif, dan netral. Data dibagi menjadi 70% data latih (2.100 tweet) dan 30% data uji (900 tweet). Eksperimen dilakukan dengan skenario yang sama untuk kedua algoritma agar hasil perbandingan bersifat adil. Sampel dataset dapat disajikan pada tabel 1

Tabel 2. Datasets dari Twitter

No	Tweet	Sentimen
1	Layanan e-government sekarang makin mudah diakses	Positif
2	Digitalisasi layanan publik sangat membantu masyarakat	Positif
3	Aplikasi pemerintah sering error dan lambat	Negatif
4	Transformasi digital ini perlu sosialisasi lebih lanjut	Netral
5	Pelayanan online pemerintah jauh lebih efisien	Positif
6	Sistem digital pemerintah sering bermasalah	Negatif
7	Saya setuju dengan pengembangan e-government	Positif
8	Keamanan data pada layanan digital masih diragukan	Negatif
9	Website pemerintah sudah cukup informatif	Netral
10	Digitalisasi membuat birokrasi lebih cepat	Positif
11	Aplikasi layanan publik sulit digunakan	Negatif
12	Transformasi digital perlu dukungan SDM	Netral
13	Layanan online pemerintah sangat membantu UMKM	Positif
14	Sistem sering down saat jam sibuk	Negatif
15	Inovasi digital pemerintah patut diapresiasi	Positif
16	Belum semua daerah siap dengan layanan digital	Netral
17	Digitalisasi meningkatkan transparansi	Positif
18	Aplikasi sering gagal login	Negatif
19	Perlu peningkatan keamanan sistem	Netral
20	Layanan digital mempermudah urusan administrasi	Positif

Preprocessing

Tahap praproses adalah langkah yang sangat penting dalam analisis sentimen karena data teks dari media sosial biasanya tidak teratur, mengandung banyak gangguan, dan memiliki variasi penulisan yang besar. Tujuan dari praproses adalah untuk membersihkan dan menstandarkan teks agar kualitas fitur yang digunakan dalam klasifikasi dapat meningkat [13]. Proses praproses yang dilakukan dalam penelitian ini mencakup pembersihan, pengubahan huruf, pemecahan kalimat, penghapusan kata umum, dan mengubah bentuk kata.

Tahap *cleansing* dilakukan untuk menghapus elemen teks yang tidak penting, seperti URL, sebutan (@username), tagar (), angka, tanda baca, emoji, serta karakter khusus lainnya. Tujuan dari proses ini adalah untuk mengurangi kebisingan dalam data dan memastikan hanya kata-kata yang memiliki arti yang disimpan untuk analisis selanjutnya. Tahap penyalarsan huruf bertujuan untuk membuat semua huruf dalam teks menjadi seragam dengan mengubah semua karakter menjadi huruf kecil. Proses ini penting untuk menghindari perbedaan makna yang disebabkan oleh variasi antara huruf besar dan kecil, sehingga kata-kata yang ejaannya sama dapat dianggap sebagai satu kesatuan yang identik. Tahap pemecahan teks dilakukan dengan memisahkan teks menjadi unit-unit kata atau token berdasarkan spasi dan tanda tertentu. Tahap penghilangan kata umum bertujuan untuk menghapus kata-kata yang sering muncul namun tidak memberikan kontribusi berarti dalam menentukan sentimen, seperti "dan", "yang", "di", dan "ke". Penghapusan kata umum ini membantu mengecilkan jumlah fitur dan meningkatkan fokus model pada kata-kata yang lebih informatif. Tahap pengembalian kata dilakukan untuk mengubah kata-kata yang memiliki imbuhan menjadi bentuk dasar. Dalam penelitian ini, pengembalian kata dalam bahasa Indonesia digunakan untuk menggabungkan variasi kata yang memiliki akar yang sama, seperti "digitalisasi", "mendigitalisasi", dan "digital" menjadi satu bentuk kata dasar. Proses ini membantu meningkatkan konsistensi fitur dan ketepatan klasifikasi..

Setelah seluruh tahapan praproses dilakukan, tahap selanjutnya adalah ekstraksi fitur menggunakan metode Term Frequency – Inverse Document Frequency (TF-IDF). TF-IDF digunakan untuk menggambarkan setiap dokumen dalam bentuk vektor angka berdasarkan seberapa penting suatu kata dalam dokumen dan seluruh koleksi teks. Bobot TF-IDF dihitung menggunakan persamaan:

$$TF - IDF_{(t,d)} = TF_{(t,d)} * IDF_{(t)} \quad (1)$$

$$\text{dengan } IDF_{(t)} = \log(N/df_{(t)}) \quad (2)$$

di mana $TF_{(t,d)}$ menyatakan frekuensi kemunculan term t pada dokumen d , N adalah jumlah total dokumen, dan $df_{(t)}$ adalah jumlah dokumen yang mengandung term t . Representasi TF-IDF ini efektif dalam menangani data teks berdimensi tinggi dan banyak digunakan dalam analisis sentimen karena mampu menekan pengaruh kata-kata umum serta menonjolkan kata-kata yang bersifat diskriminatif [14][10]

Naive Bayes

Naive Bayes adalah metode pengelompokan yang menggunakan probabilitas yang didasarkan pada Teorema Bayes, dengan anggapan bahwa semua fitur saling bebas [7]. Secara matematis, probabilitas suatu dokumen D termasuk ke dalam kelas C dihitung menggunakan persamaan berikut

$$P(C|D) = \frac{P(D|C) \cdot P(C)}{P(D)} \quad (3)$$

dimana :

$P(C|D)$: Probabilitas C dengan bukti D
 $P(C)$: Probabilitas C
 $P(D|C)$: Probabilitas D berpengaruh terhadap A
 $P(D)$: Probabilitas D

Dokumen diklasifikasikan ke dalam kelas dengan nilai probabilitas tertinggi.

Support Vector Machine (SVM)

Support Vector Machine merupakan algoritma pembelajaran terawasi yang bertujuan menemukan hyperplane terbaik yang dapat membedakan data dari berbagai kelas dengan jarak maksimum [4]. Fungsi keputusan SVM dapat dituliskan sebagai:

$$f(x) = \sum_i^n = 1a_i y_i K(x_i, x) + b \quad (4)$$

Dimana:

$f(x)$ = Fungsi untuk mengklasifikasikan data x
 n = Jumlah data yang digunakan untuk pelatihan
 a_i = Pengganda Lagrange untuk data yang ke- i
 y_i = Label untuk data latih yang ke- i
 $K(x_i, x)$ = Fungsi kernel yang digunakan untuk data latih x_i dan x
 b = Nilai bias

SVM efektif digunakan pada klasifikasi teks karena mampu menangani data berdimensi tinggi seperti vektor TF-IDF [15].

Evaluasi Model

Evaluasi model dilakukan untuk Mengevaluasi bagaimana algoritma Naive Bayes dan Support Vector Machine (SVM) bekerja dalam mengidentifikasi emosi dalam tweet dengan tepat, stabil, dan dapat digeneralisasi. Evaluasi menggunakan data uji sebanyak 900 tweet (30% dari total data) yang tidak terlibat dalam proses pelatihan, sehingga mampu merepresentasikan performa model pada data baru [16]. F1-score digunakan sebagai metrik harmonis antara presisi dan recall. Dalam penelitian ini, nilai akurasi, recall, dan F1-score dihitung untuk setiap kategori sentimen, kemudian dirata-ratakan dengan metode macro-average. Metode ini dipilih agar setiap kategori sentimen mendapatkan perhatian evaluasi yang sama dan tidak terpengaruh oleh kategori yang memiliki jumlah data lebih banyak.

Kinerja model diuji dengan menggunakan data pengujian dan dinilai memakai matriks akurasi, presisi, recall, dan F1-score. *Confusion Matrix* diterapkan untuk menghitung ukuran-ukuran tersebut. Matriks kebingungan berisi jumlah True Positive (TP), False Positive (FP), True Negative (TN), dan False Negative (FN) untuk masing-masing kategori. Dari matriks ini, dapat dihitung:

$$Akurasi = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (5)$$

$$Presisi = \frac{TP}{(TP + FP)} \quad (6)$$

menunjukkan proporsi prediksi benar positif.

$$Recall = \frac{TP}{(TP + FN)} \quad (7)$$

menunjukkan proporsi data aktual positif yang terdeteksi

$$F1\text{-score} = 2 * \frac{(\text{Presisi} \times \text{Recall})}{(\text{Presisi} + \text{Recall})} \quad (8)$$

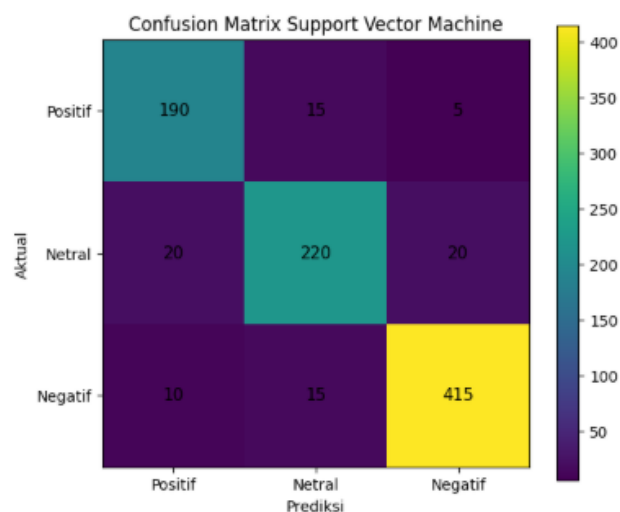
Harmonisasi presisi dan recall Penggunaan confusion matrix sangat umum dalam evaluasi klasifikasi teks. Evaluasi dilakukan untuk tiap kelas (positif, negatif, netral), sehingga dapat dibandingkan performa model dalam mengenali setiap sentimen.

3. HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil dari percobaan analisis perasaan terhadap data Twitter yang telah dikumpulkan dan diolah sebelumnya. Dataset yang dipakai dalam tahap pengujian terdiri dari 900 tweet (30% dari total 3.000 data), yang dibagi menjadi tiga kategori perasaan, yaitu positif, negatif, dan netral. Evaluasi dilakukan untuk menilai seberapa baik algoritma Naive Bayes dan Support Vector Machine (SVM) dalam mengelompokkan perasaan masyarakat mengenai topik transformasi digital yang dilakukan oleh pemerintah. Hasil evaluasi disajikan dalam bentuk confusion matrix, perhitungan metrik evaluasi, serta perbandingan akurasi antar algoritma untuk memperoleh gambaran performa yang komprehensif sebelum dilakukan pembahasan lebih lanjut.

a. Confusion Matrix

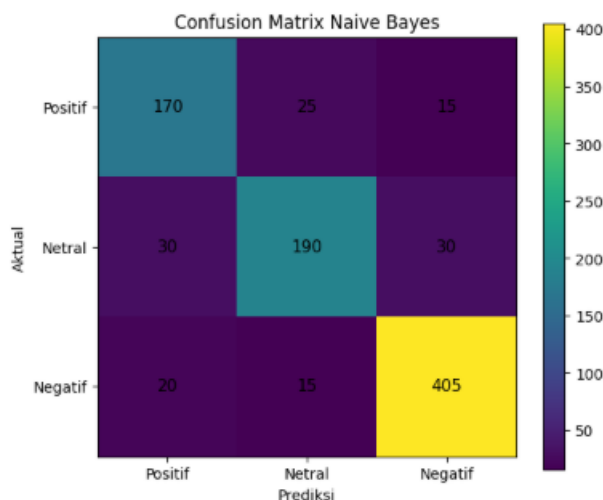
Confusion matrix adalah matriks 2×2 yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi terhadap data aktual, yang terdiri dari True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN) [17]. Confusion matrix menggambarkan kinerja masing-masing algoritma dalam mengelompokkan sentimen dari tweet menjadi tiga kategori, yaitu baik, biasa, dan buruk, berdasarkan perbandingan antara label sebenarnya dan hasil prediksi dari model. Hasil dari *confusion matrix* dari model tersebut dapat dilihat pada gambar 2 dan gambar 3 di bawah ini.



Gambar 2. Confusion Matrix SVM

Pada confusion matrix Support Vector Machine (SVM), terlihat bahwa model mampu mengklasifikasikan sebagian besar data dengan benar pada setiap kelas. Untuk kelas positif, sebanyak 190 data berhasil diprediksi dengan benar, sementara 25 data salah diklasifikasikan sebagai netral dan 5 data sebagai negatif. Pada kelas netral, SVM mengklasifikasikan 220 data secara tepat, namun masih terdapat kesalahan prediksi ke kelas positif dan negatif masing-masing sebanyak 20 data. Kinerja terbaik

SVM terlihat pada kelas negatif dengan 415 data terklasifikasi benar, menunjukkan kemampuan SVM yang kuat dalam mengenali sentimen negatif secara konsisten. Secara umum, pola ini menunjukkan bahwa kesalahan klasifikasi SVM paling sering terjadi pada kelas netral yang memiliki karakteristik linguistik cenderung ambigu.



Gambar 3. Confusion Matrix Naïve Bayes

Sementara itu, confusion matrix Naive Bayes menunjukkan jumlah prediksi benar yang juga cukup tinggi, namun masih lebih rendah dibandingkan SVM. Pada kelas positif, Naive Bayes berhasil mengklasifikasikan 170 data dengan benar, dengan kesalahan prediksi yang relatif kecil ke kelas netral dan negatif. Untuk kelas netral, sebanyak 190 data diklasifikasikan dengan tepat, namun masih terdapat kesalahan ke kelas positif dan negatif masing-masing sebanyak 30 data. Pada kelas negatif, Naive Bayes mencatat 405 prediksi benar, yang menunjukkan bahwa algoritma ini juga cukup efektif dalam mengenali sentimen negatif. Meskipun demikian, distribusi kesalahan yang lebih merata pada Naive Bayes mengindikasikan keterbatasan asumsi independensi fitur dalam membedakan sentimen yang memiliki kemiripan kata.

Secara perbandingan, kedua *confusion matrix* menunjukkan bahwa kategori netral adalah kategori yang paling sulit untuk diklasifikasikan dengan tepat. Namun, SVM menunjukkan tingkat kesalahan yang lebih sedikit dibandingkan Naive Bayes di hampir semua kategori, yang menegaskan kelebihan SVM dalam mengolah data teks dengan dimensi tinggi menggunakan representasi TF-IDF.

b. Perhitungan Manual Nilai Akurasi Metrik Evaluasi

Perhitungan metrik evaluasi dilakukan secara manual berdasarkan confusion matrix untuk memberikan gambaran transparan mengenai kinerja masing-masing algoritma. Parameter yang digunakan meliputi akurasi, presisi, recall, dan F1-score.

Akurasi dihitung sebagai perbandingan antara jumlah prediksi benar terhadap seluruh data uji.

$$\text{Akurasi} = \frac{(\text{Jumlah prediksi benar})}{(\text{Jumlah seluruh data})}$$

Untuk Naive Bayes : Jumlah prediksi benar = 170 + 190 + 405 = 765

$$\text{Akurasi} = \frac{765}{900} = 0,85 \quad 765 / 900 = 0,85$$

Untuk SVM : Jumlah prediksi benar = 190 + 210 + 415 = 815

$$\text{Akurasi} = \frac{815}{900} = 0,91$$

Berdasarkan hasil confusion matrix untuk model Naïve bayes, diperoleh:

- True Positive (TP) = 170
- False Positive (FP) = 30 + 20 = 50
- False Negative (FN) = 25 + 15 = 40

$$\text{Presisi} = \frac{TP}{(TP + FP)} = \frac{170}{(170 + 50)} = 0,77$$

$$\text{Recall} = \frac{TP}{(TP + FN)} = \frac{170}{(170 + 40)} = 0,81$$

$$\text{F1-score} = 2 * \frac{(\text{Presisi} \times \text{Recall})}{(\text{Presisi} + \text{Recall})} = 0,79$$

Perhitungan yang sama diterapkan pada kelas negatif dan netral untuk masing-masing algoritma. Nilai presisi, recall, dan F1-score kemudian dirata-ratakan menggunakan pendekatan macro-average guna memastikan setiap kelas sentimen memiliki bobot penilaian yang setara.

c. Hasil Evaluasi

Setelah dilakukan perhitungan nilai akurasi menggunakan confusion matrix, hasil yang terdapat pada table 1 dan table 2 menunjukkan keseluruhan nilai evaluasi dari algoritma naïve bayes dan support vector machine berikut

Tabel 3. Hasil Evaluasi Naive Bayes

Kelas	Presisi	Recall	F1-score
Positif	0,77	0,81	0,79
Negatif	0,83	0,76	0,79
Netral	0,90	0,92	0,91
Rata-rata	0,83	0,83	0,83

Berdasarkan tabel hasil penilaian, algoritma Naive Bayes menunjukkan performa yang cukup memuaskan di semua kategori sentimen. Pada kategori positif, nilai presisi 0,77 menunjukkan bahwa kebanyakan tweet yang diidentifikasi sebagai positif benar-benar positif, sedangkan nilai recall 0,81 menunjukkan kemampuan model yang baik dalam mendeteksi data positif yang sebenarnya. Pada kategori negatif, presisi mencapai 0,83, tetapi recall menurun menjadi 0,76, yang berarti masih ada beberapa data negatif yang salah dikategorikan ke kategori lain. Performa terbaik Naive Bayes terlihat pada kategori netral, dengan presisi 0,90, recall 0,92, dan F1-score 0,91, menunjukkan bahwa model sangat efisien dalam mengenali sentimen netral.

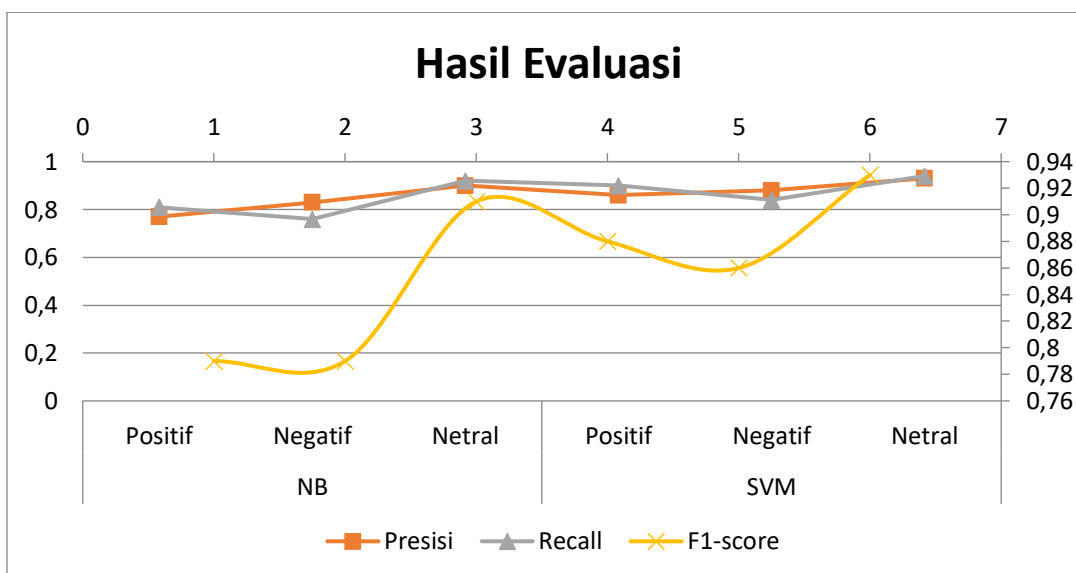
Secara keseluruhan, rata-rata F1-score Naive Bayes sebesar **0,83** menandakan performa yang stabil, meskipun masih terdapat keterbatasan dalam membedakan sentimen positif dan negatif yang memiliki kemiripan konteks.

Tabel 4. Hasil Evaluasi SVM

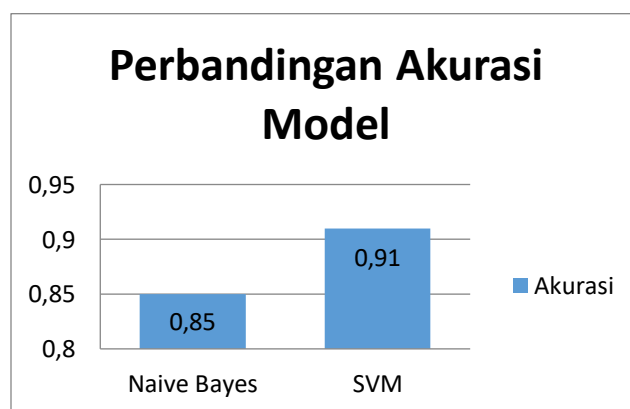
Kelas	Presisi	Recall	F1-score
Positif	0,86	0,90	0,88

Kelas	Presisi	Recall	F1-score
Negatif	0,88	0,84	0,86
Netral	0,93	0,94	0,93
Rata-rata	0,89	0,89	0,89

Hasil analisis menunjukkan bahwa Support Vector Machine (SVM) memberikan hasil yang lebih baik dibandingkan Naive Bayes di semua kategori sentimen. Dalam kategori positif, SVM memperoleh presisi 0,86 dan recall 0,90, yang menunjukkan kemampuannya yang tinggi dalam mengenali dan memprediksi sentimen positif dengan akurat. Untuk kategori negatif, presisi dan recall masing-masing mencapai angka 0,88 dan 0,84, yang menunjukkan adanya keseimbangan yang baik antara akurasi dan kelengkapan prediksi. Pada kategori netral, SVM mencapai performa puncak dengan presisi 0,93, recall 0,94, dan F1-score 0,93, yang menggarisbawahi keunggulan SVM dalam menangani sentimen yang mungkin tidak jelas. Rata-rata nilai F1-score sebesar 0,89 menunjukkan bahwa SVM lebih konsisten dan dapat dipercaya dalam mengklasifikasikan sentimen yang didasarkan pada teks TF-IDF.



Gambar 4. Hasil Evaluasi Model



Gambar 5. Grafik Perbandingan Akurasi Model

Grafik perbandingan menunjukkan bahwa SVM memiliki tingkat ketepatan yang lebih baik dibandingkan Naive Bayes. Selisih akurasi sebesar 6% mengindikasikan bahwa SVM lebih mampu membangun batas keputusan yang optimal pada data teks berdimensi tinggi.

Hasil evaluasi menunjukkan bahwa Support Vector Machine (SVM) secara konsisten mengungguli Naive Bayes pada seluruh metrik evaluasi, termasuk akurasi, presisi, recall, dan F1-score. Keunggulan ini mengindikasikan bahwa SVM lebih mampu membangun batas keputusan yang optimal pada data teks berdimensi tinggi yang direpresentasikan menggunakan TF-IDF. Temuan ini memperkuat argumen teoritis bahwa SVM lebih sesuai untuk analisis sentimen berbasis media sosial yang memiliki karakteristik data jarang dan tidak terstruktur.

Dari sudut pandang akademik, hasil penelitian ini mengonfirmasi bahwa asumsi independensi fitur pada Naive Bayes menjadi keterbatasan signifikan dalam konteks bahasa alami, khususnya ketika kata-kata dalam tweet saling berkorelasi secara semantik. Sebaliknya, SVM mampu menangkap pola kompleks antar fitur, sehingga lebih adaptif terhadap variasi ekspresi sentimen masyarakat. Temuan ini memperkaya literatur analisis sentimen berbahasa Indonesia dengan bukti empiris pada konteks kebijakan publik.

Dari perspektif kebijakan publik, dominannya kesalahan klasifikasi pada kelas netral memberikan insight penting. Sentimen netral tidak selalu menunjukkan sikap acuh, tetapi dapat merepresentasikan ketidakjelasan kebijakan, kurangnya informasi, atau fase transisi adopsi layanan digital. Oleh karena itu, pemerintah dapat memanfaatkan hasil analisis sentimen ini sebagai alat monitoring awal untuk mengidentifikasi area kebijakan yang memerlukan sosialisasi lebih lanjut atau perbaikan layanan.

Dengan demikian, penelitian ini menunjukkan bahwa analisis sentimen berbasis media sosial tidak hanya berfungsi sebagai alat teknis klasifikasi teks, tetapi juga sebagai instrumen evaluasi kebijakan berbasis data (data-driven public policy evaluation) yang relevan dalam era transformasi digital pemerintahan.

4. KESIMPULAN

Penelitian ini berhasil menerapkan metode Naive Bayes dan Support Vector Machine untuk analisis sentimen topik transformasi digital pemerintahan di Twitter. Kedua model menunjukkan kinerja sangat tinggi dengan akurasi 91% untuk SVM dan 85% untuk Naive Bayes, serta presisi dan recall yang baik untuk setiap kelas. SVM secara marginal lebih baik dalam menjaga akurasi dan mengurangi kesalahan positif/negatif dibanding NB. Hasil ini sesuai dengan literatur sebelumnya yang menyatakan SVM cenderung lebih unggul pada klasifikasi teks. Matriks kebingungan mengungkapkan sebagian kecil kesalahan dalam mengklasifikasikan tweet bermuatan kuat (negatif/positif) sebagai netral, menunjukkan area potensial untuk perbaikan. Secara keseluruhan, baik NB maupun SVM terbukti efektif sebagai algoritma analisis sentimen untuk menilai opini publik tentang transformasi digital di pemerintahan, dengan SVM menunjukkan keunggulan sedikit.

Untuk penelitian selanjutnya disarankan menggunakan data real tweet seputar transformasi digital pemerintah agar hasil lebih kontekstual. Selain itu, teknik preprocessing lanjutan seperti normalisasi bahasa gaul atau model kata khusus Bahasa Indonesia dapat meningkatkan akurasi. Penggunaan algoritma lanjutan (seperti model pembelajaran dalam berbasis word embedding atau Transformer) juga dapat dieksplorasi untuk menangani makna tersirat atau sarkasme dalam tweet. Terakhir, analisis lebih mendalam mengenai perbedaan hasil antar model (misalnya analisis fitur penting) dapat memberikan wawasan tambahan bagi penerapan kebijakan pemerintah

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dalam penulisan penelitian ini.

REFERENSI

- [1] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naïve Bayes," vol. 9, no. 2, pp. 785–795, 2022.
- [2] A. Zahra, R. Tsani, U. Budiyanto, K. Pemerintah, T. Mining, and S. V. Machine, "ANALISIS SENTIMEN PUBLIK TERHADAP KEBIJAKAN PENGIRIMAN SISWA KE BARAK MILITER MENGGUNAKAN SUPPORT VECTOR MACHINE," vol. 4, no. September, pp. 261–268, 2025.
- [3] T. P. Lestari, "Analisis Text Mining pada Sosial Media Twitter Menggunakan Metode Support Vector Machine (SVM) dan Social Network Analysis (SNA)," vol. 4, no. 3, pp. 65–71, 2022, doi: 10.37034/infef.v4i3.146.
- [4] R. Sanusi, F. D. Astuti, and I. Y. Buryadi, "Analisis sentimen pada twitter terhadap program kartu pra kerja dengan recurrent neural network," vol. 5, no. 2, pp. 89–99, 2021.
- [5] E. Fatra, T. Tri, and F. Manguma, "Analisis Sentimen Media Sosial Menggunakan Metode Naïve Bayes," vol. 2, pp. 88–98, 2025.
- [6] Z. Purwanti, "Pemodelan Text Mining untuk Analisis Sentimen Terhadap Program Makan Siang Gratis di Media Sosial X Menggunakan Algoritma Support Vector Machine (SVM)," vol. 5, no. 3, pp. 3065–3079, 2024.
- [7] I. Kurniawan, A. L. Hananto, S. S. Hilabi, and A. Hananto, "Perbandingan Algoritma Naive Bayes Dan SVM Dalam Sentimen Analisis Marketplace Pada Twitter," vol. 10, no. 1, pp. 731–740, 2023.
- [8] F. F. Haranto and B. W. Sari, "IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP," vol. 15, no. 2, pp. 171–176, 2019, doi: 10.33480/pilar.v15i2.699.
- [9] H. P. P. Zuriel and A. Fahrurrozi, "IMPLEMENTASI ALGORITMA KLASIFIKASI SUPPORT VECTOR MACHINE UNTUK ANALISA SENTIMEN PENGGUNA," pp. 149–162, 2021.
- [10] J. J. Pangaribuan and O. P. Barus, "IMPLEMENTASI ALGORITMA TF-IDF DAN SUPPORT VECTOR MACHINE TERHADAP ANALISIS PENDETEKSI KOMENTAR," vol. 13, no. 1, pp. 124–134, 2023.
- [11] S. Hidayat, H. Napitupulu, and N. Gusriani, "Penerapan Model Support Vector Machine Pada Kasus Klasifikasi Teks Berdasarkan Tujuan SDGS Ke Tiga, Empat, Dan Enam.," vol. 6, no. 2, pp. 28–37, 2024.
- [12] R. A. Fikri and R. Muliono, "Sentiment Towards Social Media Politeness Ambassadors: A Case Study Using The Naive Bayes Method," vol. 8, no. 3, pp. 127–135, 2025.
- [13] B. Hakim, "ANALISA SENTIMEN DATA TEXT PREPROCESSING PADA DATA MINING DENGAN MENGGUNAKAN MACHINE LEARNING DATA TEXT PRE-PROCESSING SENTIMENT ANALYSIS IN DATA MINING USING MACHINE LEARNING School of Computer Science and Technology , Harbin Institute of Technology," vol. 4, no. 2, pp. 16–22, 2021.
- [14] V. B. Lestari and C. A. Hutagalung, "Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM , Logistic Regression , and Naive Bayes," no. 021, pp. 36–44.
- [15] I. P. Gede *et al.*, "KLASIFIKASI JUDUL BERITA BAHASA INDONESIA MENGGUNAKAN SUPPORT VECTOR MACHINE DAN SELEKSI FITUR MUTUAL INFORMATION," vol. 22, no. 1,

pp. 69–79, 2025.

- [16] Ž. Đ. Vujović, “Classification Model Evaluation Metrics,” vol. 12, pp. 6–13, 2021, doi: 10.14569/IJACSA.2021.0120670.
- [17] F. Syahputra, D. Yani, H. Tanjung, W. Verina, and O. M. Ihsan, “Analisis Kinerja Algoritma Klasifikasi terhadap Dataset Penerimaan Pegawai Outsourcing,” vol. 14, no. April, pp. 326–335, 2025.