Performance Comparison of Naïve Bayes and SVM Methods in Sentiment Analysis of Twitter/X Users Toward Sidoarjo Regency PDAM Service Quality
[Perbandingan Kinerja Metode Naïve Bayes dan SVM dalam Analisis Sentimen Pengguna Twitter/X terhadap Kualitas Pelayanan PDAM Kabupaten Sidoarjo]

Muhammad Raihan Ardiansyah[*,1)], Hamzah Setiawan [2)], Ade Eviyanti [3)], Hindarto [4)]

1) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

2) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

3) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

4) Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

*Email Penulis Korespondensi: hamzah@umsida.ac.id

Abstract. *The growth of social media has established X (Twitter) as a vital platform for the public to express opinions on public services, including those provided by PDAM. This study aims to compare the performance of the Naïve Bayes Classifier and Support Vector Machine (SVM) methods in classifying public sentiment regarding the service quality of PDAM Sidoarjo Regency. A total of 4,339 valid tweets—comprising 571 positive, 1,706 negative, and 2,062 neutral sentiments—were collected via web scraping using Python. The data underwent preprocessing, feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF), and data balancing using the Synthetic Minority Over-sampling Technique (SMOTE) on the training set to address class distribution imbalance. Testing involved four SVM kernels (Linear, Radial Basis Function [RBF], Sigmoid, and Polynomial) and four Naïve Bayes variants (Multinomial, Complement, Bernoulli, and Gaussian) across three data-splitting scenarios. The results indicate that the SVM method consistently outperformed the Naïve Bayes method. Specifically, the SVM RBF kernel achieved the highest accuracy (75,35%), while the SVM Linear kernel achieved the highest F1-Score (74,23%). These findings demonstrate that the SVM method is more effective for multi-class sentiment classification of public opinions regarding PDAM Sidoarjo Regency's service quality and holds potential to support service quality evaluation and improvement efforts.*

Keywords; *PDAM, Naïve Bayes, SVM, Sentiment Analysis, Twitter, Public Services*

Abstrak. *Perkembangan media sosial menjadikan X (Twitter) sebagai salah satu platform penting bagi masyarakat dalam menyampaikan opini terhadap layanan publik, termasuk PDAM. Penelitian ini bertujuan membandingkan kinerja metode Naïve Bayes Classifier dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen masyarakat terhadap kualitas pelayanan PDAM Kabupaten Sidoarjo. Sebanyak 4.339 tweet valid yang terdiri atas 571 sentimen positif, 1.706 sentimen negatif, dan 2.062 sentimen netral dikumpulkan melalui teknik* scraping *menggunakan Python. Data kemudian diproses melalui tahapan* preprocessing*, ekstraksi fitur menggunakan Term Frequency-Inverse Document Frequency (TF-IDF), serta penyeimbangan data menggunakan Synthetic Minority Over-sampling Technique (SMOTE) pada data latih untuk mengatasi ketidakseimbangan distribusi kelas. Pengujian dilakukan menggunakan empat kernel SVM, yaitu Linear, Radial Basis Function (RBF), Sigmoid, dan Polynomial, serta empat varian Naïve Bayes, yaitu Multinomial, Complement, Bernoulli, dan Gaussian, pada tiga skenario pembagian data. Hasil penelitian menunjukkan bahwa metode SVM secara konsisten memberikan performa yang lebih baik dibandingkan metode Naïve Bayes. Berdasarkan hasil pengujian, SVM RBF memperoleh nilai accuracy tertinggi sebesar 75,35%, sedangkan SVM Linear memperoleh nilai F1-Score tertinggi sebesar 74,23%. Hasil tersebut menunjukkan bahwa metode SVM lebih efektif dalam melakukan klasifikasi sentimen multikelas terhadap opini masyarakat mengenai kualitas pelayanan PDAM Kabupaten Sidoarjo serta berpotensi mendukung proses evaluasi dan peningkatan kualitas layanan.*

Kata Kunci; *PDAM, Naïve Bayes, SVM, Analisis Sentimen, Twitter/X, Layanan Publik*

I. PENDAHULUAN

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

Perkembangan media sosial telah membawa perubahan pada cara masyarakat menyampaikan opini, kritik, maupun apresiasi terhadap berbagai layanan publik melalui platform digital[1]. Masyarakat kini tidak hanya menyampaikan keluhan secara langsung kepada penyedia layanan, tetapi juga mengekspresikan pengalaman, kritik, saran, maupun apresiasi melalui platform digital yang dapat diakses oleh publik secara luas. X (Twitter) menjadi salah satu platform yang banyak digunakan untuk menyampaikan kritik, keluhan, maupun apresiasi terhadap pelayanan yang diterima[2]. Karakteristik X yang bersifat terbuka dan real-time menjadikan platform ini sebagai sumber informasi yang potensial untuk mengetahui respons masyarakat terhadap suatu layanan.

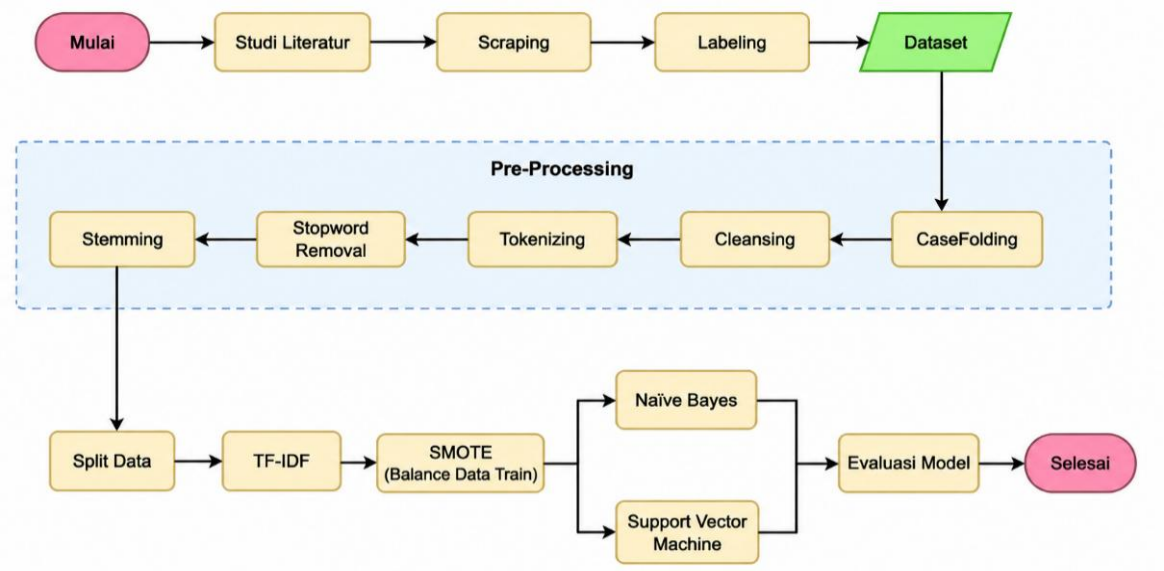
PDAM sebagai penyedia layanan air bersih juga sering menjadi topik pembahasan masyarakat di media sosial, sehingga Twitter dapat dimanfaatkan sebagai sumber data untuk memahami persepsi masyarakat terhadap kualitas pelayanan PDAM[3]. Ketersediaan air bersih merupakan kebutuhan dasar masyarakat, sehingga kualitas pelayanan PDAM menjadi aspek yang sangat penting untuk diperhatikan. Berbagai permasalahan seperti gangguan distribusi air, kualitas air yang diterima pelanggan, keterlambatan penanganan keluhan, hingga kepuasan terhadap pelayanan petugas sering kali menjadi bahan diskusi dan keluhan di media sosial. Informasi tersebut dapat dimanfaatkan oleh PDAM sebagai dasar evaluasi dalam upaya meningkatkan kualitas pelayanan [4].

Untuk memperoleh gambaran persepsi masyarakat secara sistematis, diperlukan teknik analisis yang mampu mengolah data teks dalam jumlah besar[5]. Analisis sentimen merupakan teknik yang digunakan untuk mengklasifikasikan opini masyarakat ke dalam tiga kategori, yaitu positif, netral, dan negatif. Melalui analisis sentimen, instansi pelayanan publik dapat mengetahui kecenderungan opini masyarakat serta mengidentifikasi isu-isu yang paling banyak dikeluhkan maupun diapresiasi.

Berbagai metode klasifikasi telah diterapkan dalam analisis sentimen, di antaranya Naïve Bayes Classifier dan Support Vector Machine (SVM)[6]. Kedua metode tersebut banyak digunakan karena mampu mengolah data teks dan memberikan hasil klasifikasi yang baik pada berbagai kasus analisis sentimen di media sosial[7]. Naïve Bayes dikenal memiliki proses komputasi yang sederhana dan efisien, sedangkan SVM memiliki kemampuan yang baik dalam memisahkan data pada ruang berdimensi tinggi sehingga sering menghasilkan akurasi klasifikasi yang tinggi[8].

Namun demikian, performa kedua metode tersebut dapat berbeda-beda tergantung pada karakteristik data yang digunakan, termasuk bahasa, jumlah data, serta distribusi sentimen pada media sosial[9]. Oleh karena itu, diperlukan penelitian yang secara khusus membandingkan kinerja Naïve Bayes Classifier dan Support Vector Machine (SVM) pada data sentimen pengguna X (Twitter) terkait pelayanan PDAM Kabupaten Sidoarjo. Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja Naïve Bayes Classifier dan Support Vector Machine (SVM) dalam mengklasifikasikan sentimen pengguna X (Twitter) terhadap pelayanan PDAM Kabupaten Sidoarjo. Hasil penelitian diharapkan dapat memberikan gambaran sentimen masyarakat, mengetahui metode klasifikasi yang memberikan performa terbaik, serta menjadi bahan evaluasi dalam upaya peningkatan kualitas pelayanan PDAM.

II. METODE



Gambar 1. Alur Tahapan Penelitian

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

Penelitian ini dilaksanakan melalui beberapa tahapan sistematis guna mencapai hasil yang diharapkan, sebagaimana ditampilkan pada Gambar 1.

2.1 Studi Literatur

Metode komparatif digunakan dalam penelitian ini dengan pendekatan kuantitatif karena berfokus pada pengolahan data numerik hasil klasifikasi sentimen[10]. Berbagai penelitian terdahulu telah menerapkan algoritma Naïve Bayes dan Support Vector Machine (SVM) untuk analisis sentimen pada berbagai topik, seperti ulasan aplikasi, permainan digital, hingga opini pengguna media sosial.

Kedua metode tersebut umumnya dipadukan dengan tahapan preprocessing teks dan pembobotan TF-IDF sebelum proses klasifikasi dilakukan. Hasil penelitian sebelumnya menunjukkan bahwa SVM cenderung menghasilkan performa klasifikasi yang lebih baik dibandingkan Naïve Bayes pada beberapa kasus analisis sentimen, terutama dalam menangani data teks dengan variasi kata yang tinggi.

Namun, kajian yang secara khusus membandingkan kedua metode pada sentimen masyarakat terhadap pelayanan PDAM di platform X (Twitter) masih terbatas. Oleh karena itu, penelitian ini dilakukan untuk membandingkan kinerja Naïve Bayes dan SVM dalam mengklasifikasikan sentimen pengguna terhadap pelayanan PDAM sebagai bahan evaluasi peningkatan kualitas layanan.

2.2 Scraping data

Data yang digunakan dalam penelitian ini ialah data sekunder yang berbentuk 4.349 tweet publik yang diambil dari platform X (Twitter) rentang 1 Januari 2022 hingga 20 Juni 2026. Data tersebut berisi beragam opini, keluhan, kritik, dan apresiasi masyarakat terkait layanan PDAM.

Pemilihan data dilakukan dengan mencari komentar yang memiliki kata kunci yang berhubungan dengan topik penelitian, serta hanya diambil tweet yang menggunakan bahasa Indonesia agar sesuai dengan konteks analisis sentimen yang dilakukan.

Pengumpulan data dilakukan menggunakan cara scraping memakai bahasa pemrograman Python[11]. Data dari komentar dikelompokan berdasarkan kata kunci yang berkaitan dengan layanan PDAM, seperti. PDAM, air bersih, pelayanan PDAM, Air mati, dan air PDAM tidak mengalir. Tweet yang diambil dibatasi pada bahasa Indonesia (lang:id) agar sesuai dengan konteks penelitian.

2.3 Pelabelan Data

Pada tahap pelabelan data, tweet yang telah melalui proses scraping akan diberi label secara manual ke dalam tiga kategori sentimen berawal 4349 tweet karena ada tweet yang tidak valid maka berkurang menjadi 4.339 tweet, yaitu 571 positif, 1.706 negatif, dan 2.062 netral. Proses pelabelan ini dilakukan oleh peneliti untuk memastikan kualitas dan akurasi data latih yang digunakan dalam proses klasifikasi[12].

Sentimen positif diberikan pada tweet yang berisi apresiasi atau kepuasan terhadap pelayanan PDAM, sentimen negatif diberikan pada tweet yang mengandung keluhan atau kritik, sedangkan sentimen netral diberikan pada tweet yang bersifat informatif atau tidak menunjukkan kecenderungan emosi tertentu. Dengan dilakukan pelabelan secara manual, diharapkan data yang digunakan dalam proses pelatihan model memiliki tingkat keakuratan yang lebih tinggi sehingga dapat meningkatkan performa model klasifikasi yang dihasilkan.

2.4 Preprocessing Data Teks

Pada tahapan preprocessing dilakukan pembersihan data dan juga dilakukan perapian agar teks ledih rapih[13]. Pada tahapan ini dilakuan beberapa tahapan seperti, case folding mengubah seluruh huruf menjadi huruf kecil, cleansing menghapus elemen-elemen yang tidak diperlukan, tokenizing memecah kalimat menjadi token/kata, stopword removal menghapus kata-kata yang dianggap tidak memiliki makna penting, serta stemming mengubah kata ke bentuk dasar menggunakan Sastrawi. Tahap tersebut dilakuakn bertujuan guna menghasilkan data teks yang bersih dan sudah siap untuk digunakan pada proses ekstraksi fitur.

2.5 Pembagian Data

Pada tahapan ini dilakukan pembagian dataset menjadi dua bagian, yaitu data latih dan data uji. Pada penelitian ini digunakan tiga skema pembagian data, yaitu rasio 70:30, 50:50, dan 30:70. Penggunaan tiga skema ini bertujuan untuk melihat secara komprehensif pengaruh proporsi ketersediaan data latih terhadap kinerja model klasifikasi. Data latih digunakan untuk membangun serta melatih model klasifikasi, sedangkan data uji dimanfaatkan untuk mengevaluasi kemampuan model dalam mengklasifikasikan sentimen pada tweet yang tidak digunakan selama proses pelatihan.

2.6 Pembobotan TF-IDF

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

Ekstraksi fitur teks dilakukan menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) untuk merepresentasikan data teks ke dalam bentuk numerik[14]. Pembobotan TF-IDF digunakan untuk memberikan penekanan pada kata-kata yang memiliki tingkat kepentingan lebih tinggi dalam suatu dokumen dibandingkan keseluruhan korpus. Secara konseptual, pembobotan TF-IDF dapat dirumuskan sebagai berikut:

$$TF - IDF(t, d) = TF(t, d) \times \log\left(\frac{N}{\mathrm{DF}(t)}\right) \quad (1)$$

Pada persamaan tersebut, TF(t, d) merujuk pada frekuensi kemunculan kata *t* di dalam dokumen *d*, DF(t) menunjukkan banyaknya dokumen yang memuat kata *t*, sedangkan *N* adalah jumlah keseluruhan dokumen yang terdapat dalam korpus. Dalam penelitian ini, tahap ekstraksi fitur dilakukan menggunakan metode TF-IDF yang tersedia pada library *scikit-learn*. Proses *fit* TF-IDF dilakukan hanya pada data latih, kemudian transformasi diterapkan pada data uji menggunakan *vocabulary* yang dibangun dari data latih sehingga tidak terjadi kebocoran informasi (*data leakage*). Hasil ekstraksi fitur berupa matriks numerik yang selanjutnya digunakan sebagai masukan pada proses SMOTE untuk menyeimbangkan data latih, sebelum dilakukan pelatihan dan evaluasi model Support Vector Machine (SVM) dan Naïve Bayes.

2.7 Balance Data Menggunakan SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) merupakan metode oversampling yang digunakan untuk mengatasi ketidakseimbangan jumlah data pada setiap kelas (*imbalanced data*)[15]. Ketidakseimbangan data dapat menyebabkan model klasifikasi lebih cenderung mempelajari karakteristik kelas mayoritas sehingga menurunkan kemampuan model dalam mengklasifikasikan kelas minoritas. SMOTE bekerja dengan menghasilkan data sintetis pada kelas minoritas berdasarkan kedekatan antar sampel, sehingga distribusi data latih menjadi lebih seimbang. Pada penelitian ini, SMOTE diterapkan setelah proses ekstraksi dengan fitur metode TF-IDF dan hanya pada data latih (*training data*), sedangkan data uji (*testing data*) tidak mengalami proses oversampling. Dataset hasil SMOTE selanjutnya digunakan sebagai data masukan pada proses pelatihan model Support Vector Machine (SVM) dan Naïve Bayes.

2.8 Pelatihan Model Naive Bayes

Naïve Bayes Classifier merupakan metode klasifikasi probabilistik yang didasarkan pada Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen[16]. Metode ini digunakan untuk menghitung probabilitas suatu dokumen termasuk ke dalam kelas tertentu berdasarkan kata-kata yang terkandung di dalamnya. Berikut adalah rumusnya:

$$P(H|X) = \frac{P(X|H)P(H)}{\mathrm{P}(X)} \quad (2)$$

Persamaan tersebut menegaskan bahwa penetapan kelas sentimen pada sebuah tweet ditentukan oleh dua faktor yang saling melengkapi. Nilai P(X | H) mencerminkan seberapa besar kemungkinan suatu fitur atau kata muncul dalam kelas sentimen tertentu, sementara P(H) merepresentasikan bobot awal tiap kelas berdasarkan proporsinya di dalam keseluruhan data. Adapun P(X) berperan sebagai konstanta penyeimbang yang memastikan keluaran probabilitas tetap berada dalam rentang yang valid. Dalam pelaksanaan penelitian ini, kalkulasi tersebut tidak dijalankan secara manual mengikuti persamaan di atas, melainkan diserahkan sepenuhnya kepada implementasi komputasional dengan memanfaatkan library Naïve Bayes dari paket scikit-learn. Untuk memastikan eksperimen yang mendalam, penelitian ini menguji 4 varian Naïve Bayes, yaitu: Gaussian cocok untuk data berdistribusi normal, Multinomial optimal untuk frekuensi kata diskrit teks, Bernoulli fokus pada kemunculan biner kata, dan Complement yang dirancang untuk meningkatkan performa klasifikasi pada data teks dan dikenal lebih robust terhadap ketidakseimbangan kelas dibandingkan Multinomial Naïve Bayes. Berkebalikan dengan SVM yang mengandalkan margin spasial, Naïve Bayes bekerja penuh berdasarkan teori hitung peluang (probabilitas statistik). Oleh karena itu, tantangan utamanya adalah menghadapi "kata yang tak pernah dijumpai" (zero-frequency problem). Guna mengatasi kelemahan mendasar tersebut, seluruh varian Naïve Bayes (kecuali Gaussian yang berbasis distribusi kontinu) disematkan parameter Laplace Smoothing (alpha = 1.0). Setelan tambahan berupa fit_prior = True juga dideklarasikan agar model tetap patuh pada persentase awal (prior probability) kelas yang masuk saat fase pelatihan.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

2.9 Pelatihan Model Support Vector Machine

Support Vector Machine (SVM) merupakan metode klasifikasi yang bekerja dengan mencari hyperplane optimal yang dapat memisahkan data ke dalam kelas yang berbeda[17]. SVM dikenal efektif dalam menangani data berdimensi tinggi, seperti data teks pada analisis sentimen. Metode ini berusaha memaksimalkan margin atau jarak antara data terdekat dari masing-masing kelas terhadap hyperplane. Berikut adalah rumusnya:

$$f(x) = wr^T x + b \tag{3}$$

Persamaan di atas merepresentasikan fungsi keputusan yang bekerja pada Support Vector Machine, di mana setiap vektor fitur x dipetakan ke dalam ruang berdimensi tinggi melalui kombinasi linier antara bobot w dan nilai bias b guna menetapkan posisi data relatif terhadap hyperplane pemisah. Nilai keluaran dari fungsi keputusan digunakan untuk menentukan kelas sentimen suatu data. Pada penelitian ini, klasifikasi dilakukan pada tiga kelas, yaitu positif**,** netral, dan negatif**,** dengan mekanisme multiclass yang diimplementasikan secara otomatis oleh library *scikit-learn*. Dalam konteks penelitian ini, penerapan algoritma SVM tidak dilakukan melalui perhitungan matematis secara eksplisit, melainkan diimplementasikan menggunakan library SVM dari paket *scikit-learn*. Model klasifikasi dilatih menggunakan data latih yang telah diproses melalui tahap preprocessing, penerapan Synthetic Minority Over-sampling Technique (SMOTE), serta ekstraksi fitur menggunakan Term Frequency–Inverse Document Frequency (TF-IDF) dan penyeimbangan kelas menggunakan Synthetic Minority Over-sampling Technique (SMOTE), sedangkan proses pengujian dilakukan pada data uji tanpa melalui tahap oversampling**.** Pada rumpun algoritma Support Vector Machine, fondasi pemisah kelas (hyperplane) sangat dipengaruhi oleh jenis kernel yang dipilih. Dalam studi komparatif ini, empat fungsi kernel diuji secara simultan guna melihat fleksibilitas SVM dalam menghadapi teks yang berdimensi tinggi. Nilai parameter C ditetapkan sebesar 1.0 untuk memberikan keseimbangan antara kemampuan model dalam mengklasifikasikan data latih dan kemampuan generalisasi terhadap data baru. Selain itu, parameter max_iter ditetapkan sebesar 2000 untuk membatasi jumlah iterasi pelatihan sehingga proses komputasi tetap efisien tanpa mengurangi kualitas model secara signifikan.

2.10 Evaluasi Model

Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik Accuracy, Precision, Recall, dan F1-Score untuk mengukur kinerja model klasifikasi. Tujuan dari tahapan ini ialah untuk mengetahui hasil evaluasi dari kedua algoritma, yang kemudian akan dilakukan perbandingan guna mengetahui metode mana yang memberikan hasil dan performa terbaik dalam analisis sentimen pelayanan PDAM. Untuk mengetahui kinerja model klasifikasi yang digunakan, dilakukan proses evaluasi menggunakan confusion matrix. Berdasarkan confusion matrix, diperoleh beberapa metrik evaluasi seperti accuracy, precision, recall, dan f1-score. Accuracy menunjukkan tingkat ketepatan model dalam mengklasifikasikan data secara keseluruhan.

$$\boldsymbol{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{TP + TN + FP + FN} \tag{4}$$

$$\boldsymbol{Precision} = \frac{\mathrm{TP}}{TP + FP} \tag{5}$$

$$\boldsymbol{Recall} = \frac{\mathrm{TP}}{TP + FN} \tag{6}$$

$$\boldsymbol{F1 - Score} = 2\mathrm{x}\ \frac{Precision\ x\ Recall}{Precision + Recall} \tag{7}$$

Pada penelitian ini, nilai Precision, Recall, dan F1-Score dihitung menggunakan pendekatan weighted average, sehingga kontribusi setiap kelas disesuaikan dengan jumlah data pada masing-masing kelas. Pendekatan ini dipilih karena penelitian menggunakan klasifikasi multikelas yang terdiri atas kelas positif, netral, dan negatif. Dari sisi teoritis, accuracy mencerminkan sejauh mana model mampu menghasilkan prediksi yang tepat secara agregat terhadap seluruh data uji. Precision mengukur tingkat presisi model ketika menetapkan suatu data ke dalam kelas tertentu, sementara recall menilai kapasitas model dalam menjangkau seluruh data yang sejatinya termasuk ke dalam kelas yang dimaksud. Adapun F1-score hadir sebagai representasi harmonik antara keduanya, memberikan gambaran yang lebih berimbang ketika terdapat ketimpangan antara nilai precision dan recall. Pada implementasinya, nilai Accuracy dihitung menggunakan fungsi accuracy_score, sedangkan Precision, Recall, dan F1-Score diperoleh menggunakan fungsi precision_score**,** recall_score, dan f1_score dari library *scikit-learn* dengan

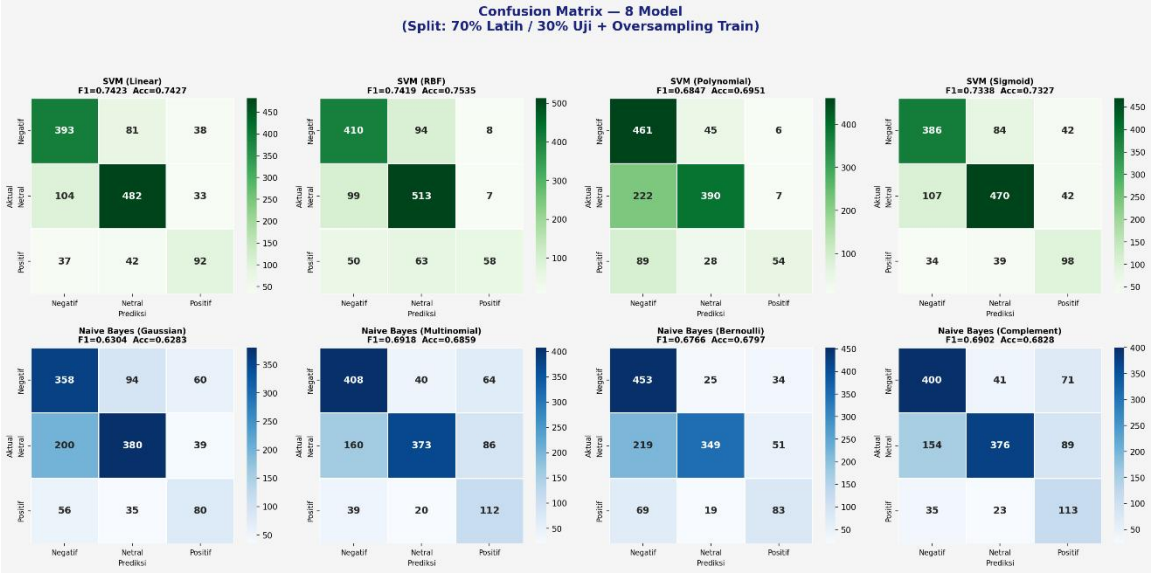
Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

parameter average = "weighted". Selain itu, *confusion matrix* digunakan untuk memberikan gambaran rinci mengenai hasil klasifikasi pada setiap kelas sehingga proses evaluasi dapat dilakukan secara lebih komprehensif.

III. HASIL DAN PEMBAHASAN

3.1 Evaluasi Model Naive Bayes dan SVM

Pengujian dilakukan terhadap delapan varian model, yaitu empat kernel Support Vector Machine (SVM) (Linear, RBF, Polynomial, dan Sigmoid) serta empat varian Naïve Bayes (Gaussian, Multinomial, Bernoulli, dan Complement). Pengujian menggunakan tiga skenario pembagian data (70%:30%, 50%:50%, dan 30%:70%), dengan data latih pada setiap skenario diseimbangkan menggunakan Synthetic Minority Over-sampling Technique (SMOTE) diterapkan sebelum proses pelatihan model.



Gambar 2. Confusion Matrix 70% Latih dan 30% Uji

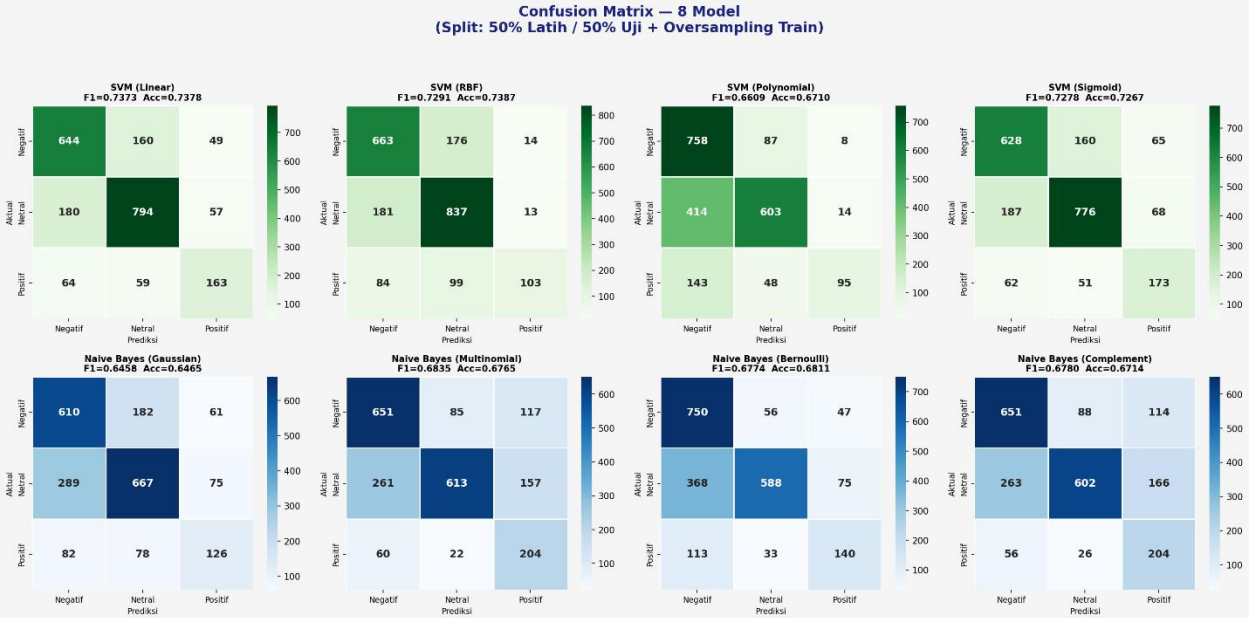
Tabel 1. Hasil evaluasi model pada skenario 70% Latih dan 30% Uji

Model	Kernel / Varian	F1-Score (Weighted)	Accuracy	Keterangan
SVM	Linear	**74,23%**	74,27%	**Terbaik F1**
SVM	RBF	74,19%	**75,35%**	**Terbaik Acc**
SVM	Sigmoid	73,38%	73,27%	-
SVM	Polynomial	68,47%	69,51%	-
Naïve Bayes	Multinomial	**69,18%**	**68,59%**	**Terbaik pada naïve bayes**
Naïve Bayes	Complement	69,02%	68,28%	-
Naïve Bayes	Bernoulli	67,66%	67,97%	-
Naïve Bayes	Gaussian	63,04%	62,83%	Terendah

Pada skenario pembagian data 70% data latih dan 30% data uji, seluruh model mampu menunjukkan performa yang baik setelah diterapkan Synthetic Minority Over-sampling Technique (SMOTE) pada bagian data latih. SVM

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

(Linear) memperoleh F1-Score tertinggi sebesar 74,23% dengan Accuracy sebesar 74,27%, sedangkan SVM (RBF) memperoleh Accuracy tertinggi sebesar 75,35% dengan F1-Score sebesar 74,19%. Sementara itu, SVM (Sigmoid) menghasilkan F1-Score sebesar 73,38% dan Accuracy sebesar 73,27%, sedangkan SVM (Polynomial) menunjukkan performa terendah pada kelompok SVM dengan F1-Score sebesar 68,47% dan Accuracy sebesar 69,51%. Pada kelompok Naïve Bayes, Multinomial Naïve Bayes memberikan performa terbaik dengan F1-Score sebesar 69,18% dan Accuracy sebesar 68,59%, diikuti oleh Complement Naïve Bayes dengan F1-Score sebesar 69,02% dan Accuracy sebesar 68,28%. Selanjutnya, Bernoulli Naïve Bayes memperoleh F1-Score sebesar 67,66% dan Accuracy sebesar 67,97%, sedangkan Gaussian Naïve Bayes menjadi varian dengan performa terendah dengan F1-Score sebesar 63,04% dan Accuracy sebesar 62,83%. Berdasarkan confusion matrix pada Gambar 2, SVM (Linear) mampu mengklasifikasikan 393 tweet negatif, 482 tweet netral, dan 92 tweet positif dengan benar. Sementara itu, SVM (RBF) memberikan hasil terbaik pada kelas netral dengan 513 prediksi benar. Pada kelompok Naïve Bayes, Gaussian Naïve Bayes menunjukkan performa terendah dengan hanya mampu mengklasifikasikan 380 tweet netral secara benar dari total 619 tweet pada kelas netral.



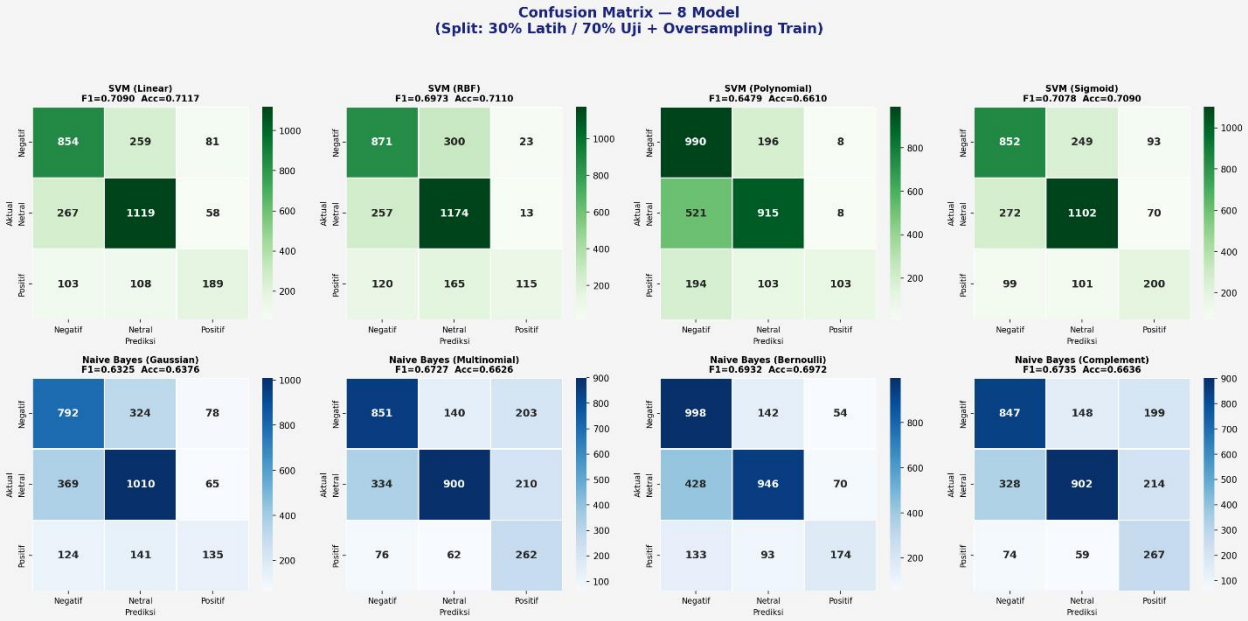
Gambar 3. Confusion Matrix 50% Latih dan 50% Uji

Tabel 2. Hasil evaluasi model pada skenario 50% Latih dan 50% Uji

Model	Kernel / Varian	F1-Score (Weighted)	Accuracy	Keterangan
SVM	Linear	**73,73%**	73,78%	**Terbaik F1**
SVM	RBF	72,91%	**73,87%**	**Terbaik Acc**
SVM	Sigmoid	72,78%	72,67%	-
SVM	Polynomial	66,09%	67,10%	-
Naïve Bayes	Multinomial	68,35%	67,65%	**Terbaik F1 (nb)**
Naïve Bayes	Complement	**67,80%**	67,14%	-
Naïve Bayes	Bernoulli	67,74%	**68,11%**	**Terbaik acc (nb)**
Naïve Bayes	Gaussian	64,58%	64,65%	Terendah

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

Pada skenario pembagian data 50% data latih dan 50% data uji, seluruh model tetap mampu menunjukkan performa yang baik setelah diterapkan Synthetic Minority Over-sampling Technique (SMOTE) pada bagian data latih saja. SVM (Linear) memperoleh F1-Score tertinggi sebesar 73,73% dengan Accuracy sebesar 73,78%, sedangkan SVM (RBF) memperoleh Accuracy tertinggi sebesar 73,87% dengan F1-Score sebesar 72,91%. Sementara itu, SVM (Sigmoid) menghasilkan F1-Score sebesar 72,78% dan Accuracy sebesar 72,67%, sedangkan SVM (Polynomial) menunjukkan performa terendah pada kelompok SVM dengan F1-Score sebesar 66,09% dan Accuracy sebesar 67,10%. Pada kelompok Naïve Bayes, Multinomial Naïve Bayes memberikan F1-Score terbaik sebesar 68,35%, sedangkan Bernoulli Naïve Bayes memperoleh Accuracy tertinggi sebesar 68,11%. Selanjutnya, Complement Naïve Bayes memperoleh F1-Score sebesar 67,80% dan Accuracy sebesar 67,14%, sedangkan Gaussian Naïve Bayes menjadi varian dengan performa terendah dengan F1-Score sebesar 64,58% dan Accuracy sebesar 64,65%. Berdasarkan confusion matrix pada Gambar 4, SVM (Linear) mampu mengklasifikasikan 644 tweet negatif, 794 tweet netral, dan 163 tweet positif dengan benar. Sementara itu, SVM (RBF) memberikan hasil terbaik pada kelas netral dengan 837 prediksi benar. Pada kelompok Naïve Bayes, Gaussian Naïve Bayes menunjukkan performa terendah dengan hanya mampu mengklasifikasikan 667 tweet netral secara benar dari total 1.031 tweet pada kelas netral.



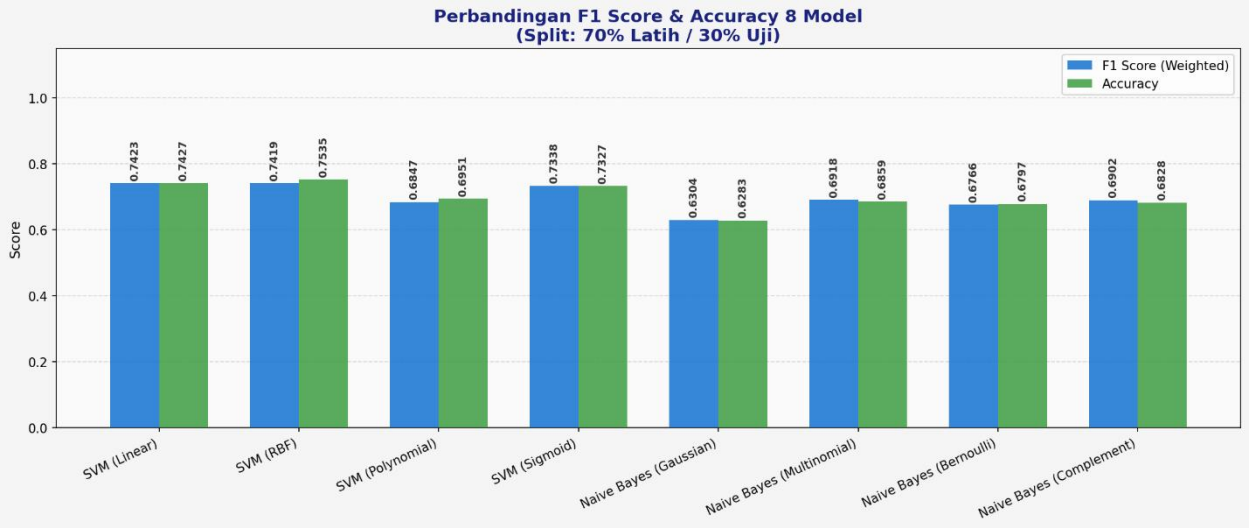
Gambar 4. Confusion Matrix 30% Latih dan 70% Uji

Tabel 3. Hasil evaluasi model pada skenario 30% Latih dan 70% Uji

Model	Kernel / Varian	F1-Score (Weighted)	Accuracy	Keterangan
SVM	Linear	**70,90%**	**71,17%**	**Terbaik F1 dan acc**
SVM	RBF	69,73%	71,10%	-
SVM	Sigmoid	70,78%	70,90%	-
SVM	Polynomial	64,79%	66,10%	-
Naïve Bayes	Multinomial	67,27%	66,26%	-
Naïve Bayes	Complement	67,35%	66,36%	-
Naïve Bayes	Bernoulli	**69,32%**	**69,72%**	**Terbaik F1 dan acc (nb)**
Naïve Bayes	Gaussian	63,25%	63,76%	Terendah

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

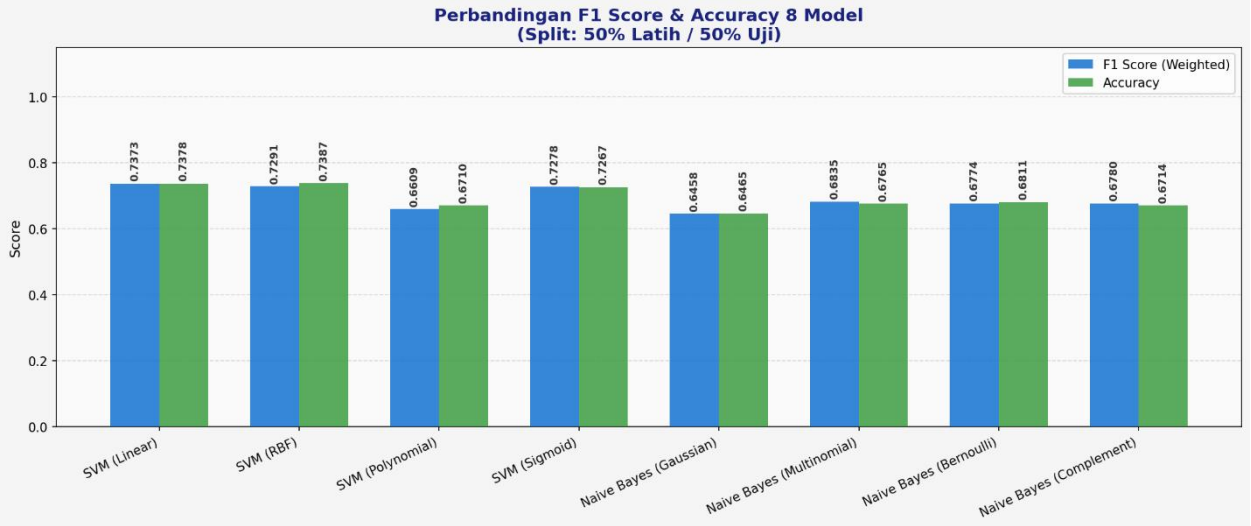
Pada skenario pembagian data 30% data latih dan 70% data uji, seluruh model tetap mampu memberikan hasil klasifikasi yang baik setelah diterapkan Synthetic Minority Over-sampling Technique (SMOTE) pada bagian data latih saja. SVM (Linear) memperoleh F1-Score tertinggi sebesar 70,90% sekaligus Accuracy tertinggi sebesar 71,17%. SVM (RBF) menghasilkan F1-Score sebesar 69,73% dan Accuracy sebesar 71,10%, sedangkan SVM (Sigmoid) memperoleh F1-Score sebesar 70,78% dan Accuracy sebesar 70,90%. SVM (Polynomial) menunjukkan performa terendah pada kelompok SVM dengan F1-Score sebesar 64,79% dan Accuracy sebesar 66,10%. Pada kelompok Naïve Bayes, Bernoulli Naïve Bayes memberikan performa terbaik dengan F1-Score sebesar 69,32% dan Accuracy sebesar 69,72%, diikuti oleh Complement Naïve Bayes dengan F1-Score sebesar 67,35% dan Accuracy sebesar 66,36%. Selanjutnya, Multinomial Naïve Bayes memperoleh F1-Score sebesar 67,27% dan Accuracy sebesar 66,26%, sedangkan Gaussian Naïve Bayes menjadi varian dengan performa terendah dengan F1-Score sebesar 63,25% dan Accuracy sebesar 63,76%. Berdasarkan confusion matrix pada Gambar 6, SVM (Linear) mampu mengklasifikasikan 854 tweet negatif, 1.119 tweet netral, dan 189 tweet positif dengan benar. Sementara itu, Bernoulli Naïve Bayes sebagai model terbaik pada kelompok Naïve Bayes mampu mengklasifikasikan 998 tweet negatif, 946 tweet netral, dan 174 tweet positif dengan benar, sedangkan Gaussian Naïve Bayes menunjukkan performa terendah dengan hanya mampu mengklasifikasikan 792 tweet negatif, 1.010 tweet netral, dan 135 tweet positif secara benar.



Gambar 5. Perbandingan F1 Score Accuracy 70% Latih dan 30% Uji

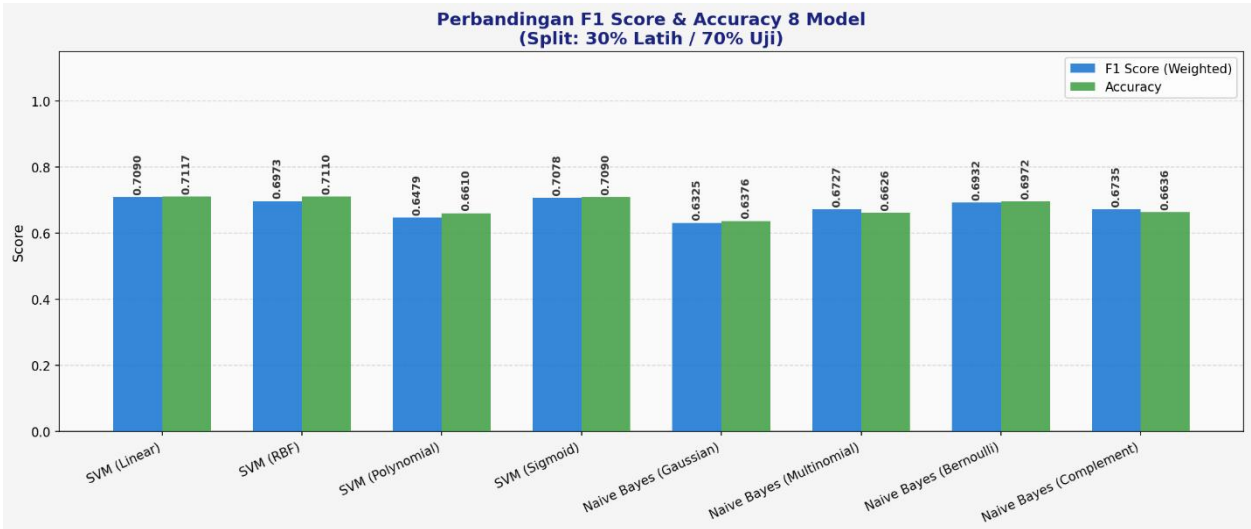
Berdasarkan Gambar 5, terlihat perbandingan nilai F1-Score (Weighted) dan Accuracy dari delapan model yang diuji pada skenario pembagian data 70% data latih dan 30% data uji setelah penerapan SMOTE pada data latih. Secara umum, seluruh model SVM menunjukkan kinerja yang lebih baik dibandingkan varian Naïve Bayes. SVM Linear memperoleh F1-Score tertinggi sebesar 74,23%, sedangkan Accuracy tertinggi dicapai oleh SVM RBF sebesar 75,35%. Pada kelompok SVM, kernel Polynomial menghasilkan performa terendah dengan F1-Score 68,47% dan Accuracy 69,51%. Sementara itu, pada kelompok Naïve Bayes, Multinomial Naïve Bayes memberikan F1-Score tertinggi sebesar 69,18%, sedangkan Accuracy tertinggi diperoleh Bernoulli Naïve Bayes sebesar 67,97%. Gaussian Naïve Bayes menjadi model dengan performa terendah pada kelompok Naïve Bayes dengan F1-Score 63,04% dan Accuracy 62,83%. Hasil tersebut menunjukkan bahwa pada skenario 70% data latih dan 30% data uji, model SVM, khususnya kernel Linear dan RBF, mampu memberikan performa klasifikasi yang lebih baik dibandingkan seluruh varian Naïve Bayes.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.



Gambar 6. Perbandingan F1 Score Accuracy 50% Latih dan 50% Uji

Berdasarkan Gambar 6, terlihat perbandingan nilai F1-Score (Weighted) dan Accuracy dari delapan model yang diuji pada skenario pembagian data 50% data latih dan 50% data uji setelah penerapan SMOTE pada data latih. Secara umum, model SVM masih menunjukkan performa yang lebih baik dibandingkan varian Naïve Bayes. SVM Linear memperoleh F1-Score tertinggi sebesar 73,73%, sedangkan Accuracy tertinggi dicapai oleh SVM RBF sebesar 73,87%. Pada kelompok SVM, kernel Polynomial menghasilkan performa terendah dengan F1-Score sebesar 66,09% dan Accuracy sebesar 67,10%. Sementara itu, pada kelompok Naïve Bayes, Multinomial Naïve Bayes memperoleh F1-Score tertinggi sebesar 68,35%, sedangkan Accuracy tertinggi dicapai oleh Bernoulli Naïve Bayes sebesar 68,11%. Gaussian Naïve Bayes menjadi model dengan performa terendah pada kelompok Naïve Bayes dengan F1-Score sebesar 64,58% dan Accuracy sebesar 64,65%. Hasil tersebut menunjukkan bahwa pada skenario pembagian data 50% data latih dan 50% data uji, model SVM tetap memberikan performa klasifikasi yang lebih baik dibandingkan seluruh varian Naïve Bayes, meskipun terjadi sedikit penurunan nilai F1-Score dan Accuracy dibandingkan skenario 70% data latih dan 30% data uji.



Gambar 7. Perbandingan F1 Score Accuracy 30% Latih dan 70% Uji

Berdasarkan Gambar 7, terlihat perbandingan nilai F1-Score (Weighted) dan Accuracy dari delapan model yang diuji pada skenario pembagian data 30% data latih dan 70% data uji setelah penerapan SMOTE pada data latih. Secara umum, model SVM masih menunjukkan performa yang bahkan lebih baik jika dibandingkan varian Naïve Bayes. SVM Linear memperoleh F1-Score tertinggi sebesar 70,90% sekaligus Accuracy tertinggi sebesar 71,17%.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

Pada kelompok SVM, kernel Polynomial menghasilkan performa terendah dengan F1-Score sebesar 64,79% dan Accuracy sebesar 66,10%. Sementara itu, pada kelompok Naïve Bayes, Bernoulli Naïve Bayes memperoleh F1-Score tertinggi sebesar 69,32% sekaligus Accuracy tertinggi sebesar 69,72%, sedangkan Gaussian Naïve Bayes menjadi model dengan performa terendah dengan F1-Score sebesar 63,25% dan Accuracy sebesar 63,76%. Hasil tersebut menunjukkan bahwa meskipun proporsi data uji lebih besar dibandingkan data latih, model SVM tetap memberikan performa klasifikasi yang lebih baik dibandingkan seluruh varian Naïve Bayes. Namun, secara keseluruhan nilai F1-Score dan Accuracy pada skenario ini mengalami sedikit penurunan dibandingkan skenario 70% data latih dan 30% data uji, yang menunjukkan bahwa berkurangnya jumlah data latih dapat memengaruhi kemampuan model dalam melakukan klasifikasi secara optimal.

Tabel 4 berikut merangkum perbandingan lengkap seluruh hasil evaluasi keenam belas konfigurasi pengujian berdasarkan F1-Score dan Accuracy tertinggi per skenario:

Tabel 4. Rekap Perbandingan Kinerja Seluruh Model di Tiga Skenario

Model	Skenario	Train %	Test %	F1-Score	Accuracy
SVM Linear	70/30	70	30	**74,23%**	74,27%
SVM RBF	70/30	70	30	74,19%	**75,35%**
SVM Sigmoid	70/30	70	30	73,38%	73,27%
SVM Polynomial	70/30	70	30	68,47%	69,51%
SVM Linear	50/50	50	50	**73,73%**	73,78%
SVM RBF	50/50	50	50	72,91%	**73,87%**
SVM Sigmoid	50/50	50	50	72,78%	72,67%
SVM Polynomial	50/50	50	50	66,09%	67,10%
SVM Linear	30/70	30	70	**70,90%**	**71,17%**
SVM RBF	30/70	30	70	69,73%	71,10%
SVM Sigmoid	30/70	30	70	70,78%	70,90%
SVM Polynomial	30/70	30	70	64,79%	66,10%
NB Multinomial	70/30	70	30	69,18%	**68,59%**
NB Gaussian	70/30	70	30	63,04%	62,83%
NB Bernoulli	70/30	70	30	67,66%	**67,97%**
NB Complement	70/30	70	30	69,02%	68,28%
NB Multinominal	50/50	50	50	**68,35%**	67,65%
NB Gaussian	50/50	50	50	64,58%	64,65%
NB Bernoulli	50/50	50	50	67,74%	**68,11%**
NB Complement	50/50	50	50	67,80%	67,14%
NB Multinominal	30/70	30	70	67,27%	66,26%
NB Gaussian	30/70	30	70	63,25%	63,76%
NB Bernoulli	30/70	30	70	**69,32%**	**69,72%**
NB Complement	30/70	30	70	67,35%	66,36%

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

Berdasarkan Tabel 4, dapat diketahui bahwa performa seluruh model dipengaruhi oleh proporsi pembagian data latih dan data uji. Secara umum, skenario dengan komposisi 70% data latih dan 30% data uji memberikan hasil terbaik dibandingkan skenario 50%:50% dan 30%:70%. Pada algoritma Support Vector Machine (SVM), kernel RBF memperoleh nilai Accuracy tertinggi sebesar 75,35%, sedangkan kernel Linear menghasilkan F1-Score tertinggi sebesar 74,23%. Seiring berkurangnya proporsi data latih, nilai F1-Score dan Accuracy pada seluruh kernel SVM cenderung menurun, menunjukkan bahwa jumlah data latih yang lebih besar mampu meningkatkan kemampuan model dalam mempelajari pola data. Pada algoritma Naïve Bayes, varian Multinomial menghasilkan F1-Score tertinggi pada skenario 70%:30% dan 50%:50%, sedangkan varian Bernoulli memperoleh Accuracy tertinggi pada kedua skenario tersebut dan menjadi model terbaik pada skenario 30%:70% dengan F1-Score sebesar 69,32% serta Accuracy sebesar 69,72%. Sebaliknya, Gaussian Naïve Bayes secara konsisten memberikan performa terendah pada seluruh skenario pengujian. Secara keseluruhan, hasil rekapitulasi menunjukkan bahwa algoritma SVM memiliki kinerja yang lebih baik dibandingkan Naïve Bayes, serta penggunaan proporsi data latih yang lebih besar menghasilkan performa klasifikasi yang lebih optimal.

3.2 Analisis WordCloud



(a) (b) (c)

Gambar 8. WordCloud Negatif Netral Positif

Berdasarkan wordcloud Gambar 8 (a) pada kelas negatif, terlihat bahwa kata-kata yang paling sering muncul adalah mati, air, daerah, rumah, layanan, distribusi, tolong, dan bersih. Kata “mati” menjadi kata yang paling dominan, yang menunjukkan banyak pengguna mengeluhkan gangguan atau tidak mengalirnya air PDAM. Selain itu, munculnya kata seperti “tolong”, “info”, dan “layak” menunjukkan adanya keluhan serta harapan masyarakat terhadap peningkatan kualitas layanan distribusi air. Wordcloud ini menggambarkan bahwa sentimen negatif didominasi oleh permasalahan distribusi, kualitas air, dan gangguan layanan PDAM di beberapa daerah.

Berdasarkan wordcloud Gambar 8 (b) pada kelas netral, terlihat bahwa kata-kata yang paling sering muncul adalah air, bersih, mati, keruh, langgan, rumah, daerah, dan tagih. Kemunculan kata-kata tersebut menunjukkan bahwa sebagian besar tweet netral membahas informasi atau kondisi layanan air PDAM tanpa memberikan opini yang terlalu positif maupun negatif. Selain itu, terdapat kata seperti filter, bayar, listrik, dan lokasi yang menunjukkan pembahasan terkait penggunaan air, pembayaran tagihan, maupun kondisi distribusi di wilayah tertentu. Wordcloud ini menggambarkan bahwa sentimen netral didominasi oleh percakapan informatif mengenai layanan dan distribusi air PDAM.

Berdasarkan wordcloud Gambar 8 (c) pada kelas positif, terlihat bahwa kata-kata yang paling sering muncul adalah alir, bersih, terima, delta, rumah, layanan, masyarakat, dan distribusi. Kata “terima” dan “bersih” menunjukkan adanya apresiasi masyarakat terhadap kualitas air dan pelayanan PDAM yang dianggap baik. Selain itu, munculnya kata seperti bantu, cepat, dan respon menggambarkan adanya kepuasan pengguna terhadap penanganan layanan dan distribusi air. Wordcloud ini menunjukkan bahwa sentimen positif didominasi oleh tanggapan masyarakat yang merasa terbantu dengan layanan, distribusi air yang lancar, serta kualitas air yang bersih dan layak digunakan.

VII. SIMPULAN

Penelitian ini berhasil mengklasifikasikan sentimen masyarakat terhadap pelayanan PDAM menggunakan 4.339 tweet yang terdiri atas 571 sentimen positif, 1.706 sentimen negatif, dan 2.062 sentimen netral. Pengujian dilakukan terhadap delapan model, yaitu empat kernel Support Vector Machine (SVM) dan empat varian Naïve Bayes, pada tiga skenario pembagian data (70%:30%, 50%:50%, dan 30%:70%) dengan menerapkan Synthetic Minority Over-sampling Technique (SMOTE) pada data latih. Hasil penelitian menunjukkan bahwa metode SVM secara konsisten memberikan kinerja yang lebih baik dibandingkan Naïve Bayes pada seluruh skenario pengujian. Performa terbaik diperoleh SVM Linear dengan F1-Score sebesar 74,23%, sedangkan Accuracy tertinggi dicapai SVM RBF sebesar 75,35% pada skenario 70% data latih dan 30% data uji. Pada kelompok Naïve Bayes,

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..

Multinomial memberikan performa terbaik pada skenario 70%:30% dan 50%:50%, sedangkan Bernoulli menjadi varian terbaik pada skenario 30%:70%. Hasil pengujian juga menunjukkan bahwa semakin besar proporsi data latih, semakin baik performa model yang dihasilkan. Analisis WordCloud memperlihatkan bahwa sentimen negatif didominasi oleh keluhan mengenai gangguan distribusi dan kualitas air, sentimen netral berisi informasi terkait layanan PDAM, sedangkan sentimen positif lebih didominasi oleh komentar berisi apresiasi terhadap kualitas air dan respons pelayanan. Untuk penelitian selanjutnya disarankan menggunakan dataset yang lebih besar, mengeksplorasi teknik penyeimbangan data selain SMOTE, serta membandingkan performa model dengan metode berbasis deep learning seperti IndoBERT atau LSTM.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada seluruh pihak yang telah memberikan dukungan, bimbingan, dan bantuan selama proses penyelesaian penelitian ini, baik dalam aspek akademik maupun nonakademik. Arahan ilmiah, saran yang membangun, serta penyediaan fasilitas pendukung telah memberikan kontribusi yang berarti sehingga penelitian ini dapat diselesaikan dengan baik. Selain itu, dukungan moral dan motivasi yang diberikan selama proses penelitian turut berperan dalam menjaga kelancaran dan kualitas penyusunan karya ilmiah ini. Penulis berharap penelitian ini dapat memberikan manfaat bagi pembaca serta menjadi referensi bagi penelitian selanjutnya.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

REFERENSI

[1] A. Septiana *et al.*, “Dalam Analisis Sentimen Pengguna Twitter X Terhadap,” vol. 8, no. 1, pp. 323–330, 2024.

[2] F. F. Mailoa, “Analisis sentimen data twitter menggunakan metode text mining tentang masalah obesitas di indonesia,” *J. Inf. Syst. Public Heal.*, vol. 6, no. 1, p. 44, 2021, doi: 10.22146/jisph.44455.

[3] P. Covid- and M. A. Pono, “Persepsi Masyarakat Terhadap aktivitas Publikasi Perumda air Minum Kota Makassar Dalam Menangani Keluhan Pelanggan Dimasa,” pp. 78–89.

[4] H. Zahra, S. Santoso, and I. A. Kurniawan, “Penerapan sistem informasi pelayanan publik pada aplikasi simpel tkr pdam tirta kerta raharja kabupaten tangerang,” vol. 7, no. September, pp. 87–98, 2024.

[5] F. Ali, A. Setiawan, R. Nuraini, and S. Fathonah, “TINJAUAN SISTEMATIS LITERATUR : ANALISIS SENTIMEN TERHADAP,” vol. 9, no. 5, pp. 7995–8003, 2025.

[6] F. M. Sarimole, “Analisis Sentimen Terhadap Aplikasi Satu Sehat Pada Twitter Menggunakan Algoritma Naive Bayes Dan Support Vector Machine,” vol. 5, no. 3, pp. 783–790, 2024.

[7] R. Mursyid, A. D. Indriyanti, J. Teknik, I. Fakultas, T. Universitas, and N. Surabaya, “Perbandingan Akurasi Metode Analisis Sentimen Untuk Evaluasi Opini Pengguna Pada Platform Media Sosial (Studi Kasus : Twitter),” vol. 06, pp. 371–383, 2024.

[8] M. Difanada, H. Setiawan, and I. A. Kautsar, “Penerapan metode support vector machine (svm) untuk memprediksi pemilihan karir bagi alumni umsida,” vol. 10, no. 4, pp. 2906–2916, 2025.

[9] J. R. Informasi *et al.*, “ANALISIS KOMPREHENSIF KINERJA MODEL KLASIFIKASI SENTIMEN : EVALUASI LINTAS METRIK PADA DATASET TWEET FILM BAHASA,” vol. 14, no. 1, 2025.

[10] L. A. Qomariah *et al.*, “KLASIFIKASI SENTIMEN ULASAN MINECRAFT PADA GOOGLE PLAY STORE : STUDI KOMPARATIF NAÏVE BAYES DAN RANDOM FOREST,” vol. 10, no. 1, pp. 1227–1234, 2026.

[11] A. Rahman and H. Suroyo, “Analisis Data Produk Elektronik Di E-Commerce Dengan Metode Algoritma K-Means Menggunakan Python,” vol. 3, no. 2, pp. 11–18, 2021.

[12] Y. Asri, W. N. Suliyanti, D. Kuswardani, and M. Fajri, “Pelabelan Otomatis Lexicon Vader dan Klasifikasi Naive Bayes dalam menganalisis sentimen data ulasan PLN Mobile,” vol. 15, no. 2, pp. 264–275, 2022.

[13] U. Khairani *et al.*, “PENGARUH TAHAPAN PREPROCESSING TERHADAP MODEL INDOBERT DAN THE INFLUENCE OF PREPROCESSING STAGES ON INDOBERT AND INDOBERTWEET MODELS FOR EMOTION DETECTION IN INSTAGRAM NEWS,” vol. 11, no. 4, 2024, doi: 10.25126/jtiik.1148315.

[14] A. Firizkiansah, A. Muhammad, I. R. Maulana, U. S. Indonesia, and K. Bekasi, “Optimasi Klasifikasi Data Teks Menggunakan Algoritma Logistic Regression dengan TF-IDF dan SMOTE,” vol. 2, no. 1, pp. 29–36, 2025.

[15] M. Sulistiyono *et al.*, “Implementasi Algoritma Synthetic Minority Over - Sampling Technique untuk Menangani Ketidakseimbangan Kelas pada Dataset Klasifikasi,” vol. 10, pp. 445–459, 2021.

[16] E. Martantoh, N. Yanih, T. Informatika, U. Panca, and S. Bekasi, “Implementasi Metode Naïve Bayes Untuk Klasifikasi Karakteristik Kepribadiaan Siswa Di Sekolah MTS Darussa ’ adah Menggunakan PHP MySQL,” vol. 3, no. 2, pp. 166–175, 2022.

[17] D. R. Nurqotimah, A. N. Khudori, and R. S. Pradini, “JOURNAL OF APPLIED COMPUTER SCIENCE AND TECHNOLOGY (JACOST) Implementasi Algoritma Support Vector Machine (SVM) Untuk Klasifikasi Penyakit Stroke,” vol. 5, no. 2, pp. 179–185, 2024.

Conflict of Interest Statement:
The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.
Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted..