UNIVERSITAS
MUHAMMADIYAH
SIDOARJO

UMSIDA
DARI SINI PENCERAHAN BERSEMI

Performance Comparison of Naïve Bayes and SVM Methods in Sentiment Analysis of Twitter/X Users Toward Sidoarjo Regency PDAM Service Quality

Oleh :

Muhammad Raihan Ardiansyah

Dosen Pembimbing:

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

Dosen Penguji 1:

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

Dosen Penguji 2:

Dr. Hindarto, S.Kom., MT.

Program Studi Informatika

Universitas Muhammadiyah Sidoarjo

2026

Latar belakang

Perkembangan media sosial, khususnya X (Twitter), telah mengubah cara masyarakat menyampaikan opini terhadap layanan publik, termasuk pelayanan PDAM Kabupaten Sidoarjo. Banyaknya tweet yang berisi keluhan, kritik, maupun apresiasi menjadikan media sosial sebagai sumber data yang potensial untuk dianalisis. Oleh karena itu, penelitian ini membandingkan metode Naïve Bayes dan Support Vector Machine (SVM) untuk menentukan metode yang paling efektif dalam analisis sentimen terhadap pelayanan PDAM Kabupaten Sidoarjo.

Rumusan masalah

1. Bagaimana proses pengumpulan dan pengolahan data tweet yang berkaitan dengan pelayanan PDAM pada platform X (Twitter)?
2. Bagaimana implementasi metode Naïve Bayes Classifier dan Support Vector Machine (SVM) guna mengklasifikasikan sentimen tweet masyrakat pengguna media sosial pada pelayanan PDAM?
3. Bagaimana komparasi kinerja metode Naïve Bayes dan Support Vector Machine (SVM) guna melakukan analisis sentimen pelayanan PDAM berdasarkan metrik evaluasi akurasi, *precision*, *recall*, dan *f1-score*?
4. Metode klasifikasi manakah yang menunjukkan performa paling optimal dalam mengklasifikasikan sentimen pengguna Twitter/X terhadap kualitas pelayanan PDAM?

Dataset



berjumlah 4.349 data tweet pada twitter setelah validasi pembersihan tweet yang valid 4.339 data, rentang pengambilan:

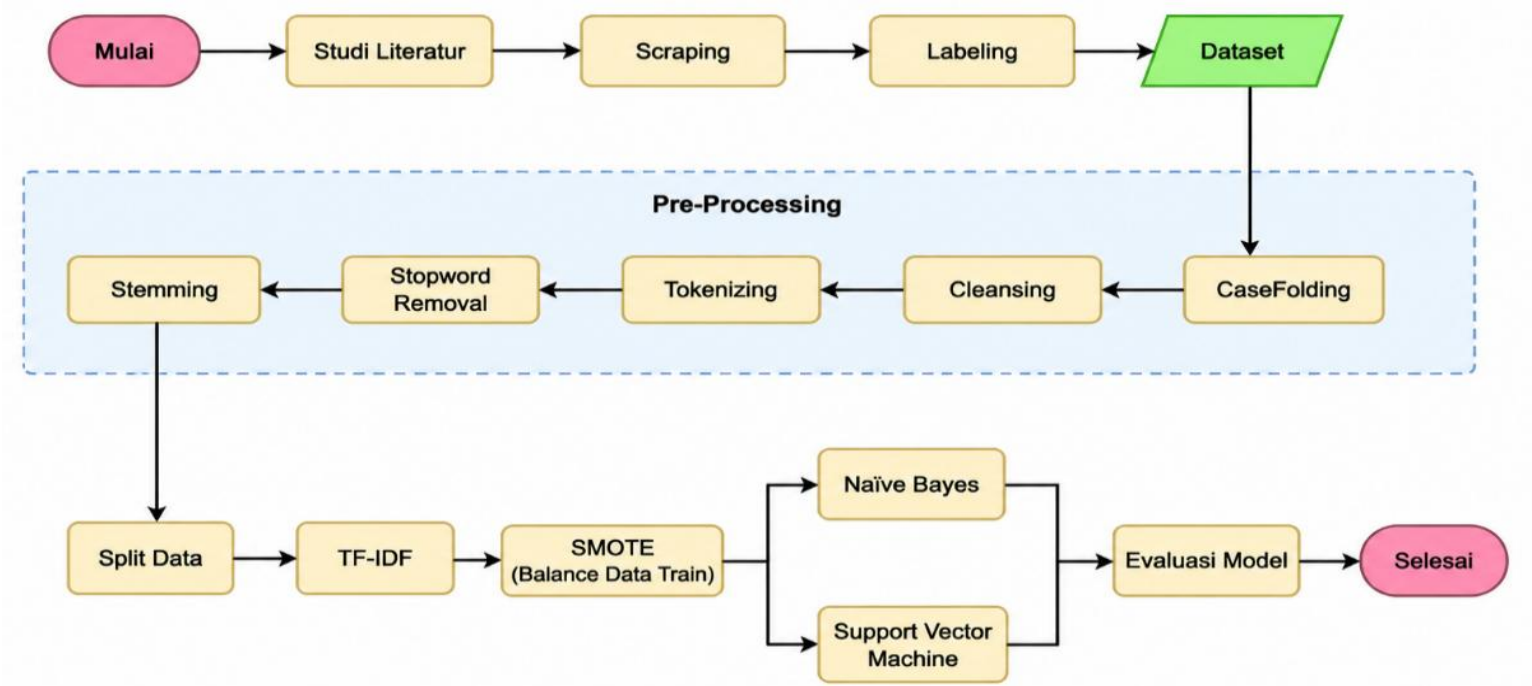
1 Januari 2022 – 20 Juni 2026

Distribusi Sentimen

Total 4.339 data dan dilakukan labelling manual yang memiliki 571 Positif, 1706 Negatif, 2062 Netral terlabel secara manual, kemudian dilakukanya preprocessing melalui case folding, cleansing, tokenizing, stopword removal, stemming, kemudian split data, kemudian pembobotan ke numerik guna TF-IDF (Term Frequency-Inverse Document Frequency), dan perlatihan model 4 varian Naïve Bayes dan 4 kernel SVM, serta evaluasi dan perbandingan Accuracy, F1-Score, Confusion Matrix.

Alur penelitian

Penelitian ini menggunakan pendekatan kuantitatif komparatif dengan data sebanyak 4.339 tweet pengguna X (Twitter) terkait pelayanan PDAM Kabupaten Sidoarjo. Data melalui tahap preprocessing, kemudian dibagi ke dalam tiga skenario, yaitu 70:30, 50:50, dan 30:70. Selanjutnya, data latih diseimbangkan menggunakan SMOTE, dilakukan ekstraksi fitur menggunakan TF-IDF, kemudian diklasifikasikan menggunakan empat kernel Support Vector Machine (SVM) dan empat varian Naïve Bayes. Kinerja setiap model dievaluasi berdasarkan Accuracy, Precision, Recall, F1-Score, dan confusion matrix.



Split Data

• Dataset dibagi menjadi data latih dan data uji menggunakan tiga skema, yaitu 70:30, 50:50, dan 30:70.
• Tujuan penggunaan tiga skema adalah untuk menganalisis pengaruh proporsi data latih terhadap performa model klasifikasi.
• Data latih digunakan untuk membangun model, sedangkan data uji digunakan untuk mengevaluasi kemampuan model pada data yang belum pernah dilihat sebelumnya.
• Untuk mengatasi ketidakseimbangan kelas (imbalanced data), diterapkan teknik Oversampling berbasis Synthetic Minority Over-sampling Technique (SMOTE).
• SMOTE bekerja dengan menghasilkan data sintetis pada kelas minoritas berdasarkan kedekatan antar sampel, sehingga distribusi data latih menjadi lebih seimbang.
• Oversampling hanya diterapkan pada data latih setelah proses splitting, sedangkan data uji tetap dibiarkan asli.
• Langkah ini bertujuan menjaga keseimbangan distribusi kelas selama pelatihan dan mencegah terjadinya data leakage pada proses pengujian.

Split Data

Tabel 2. Skenario 70% Latih dan 30% Uji

Kategori	Latih (Sebelum Oversample)	Latih (Sesudah Oversample)	Uji (Test)
Negatif	1194	1194	512
Netral	1443	1443	619
Positif	400	400	171
Total	3037	3037	1302

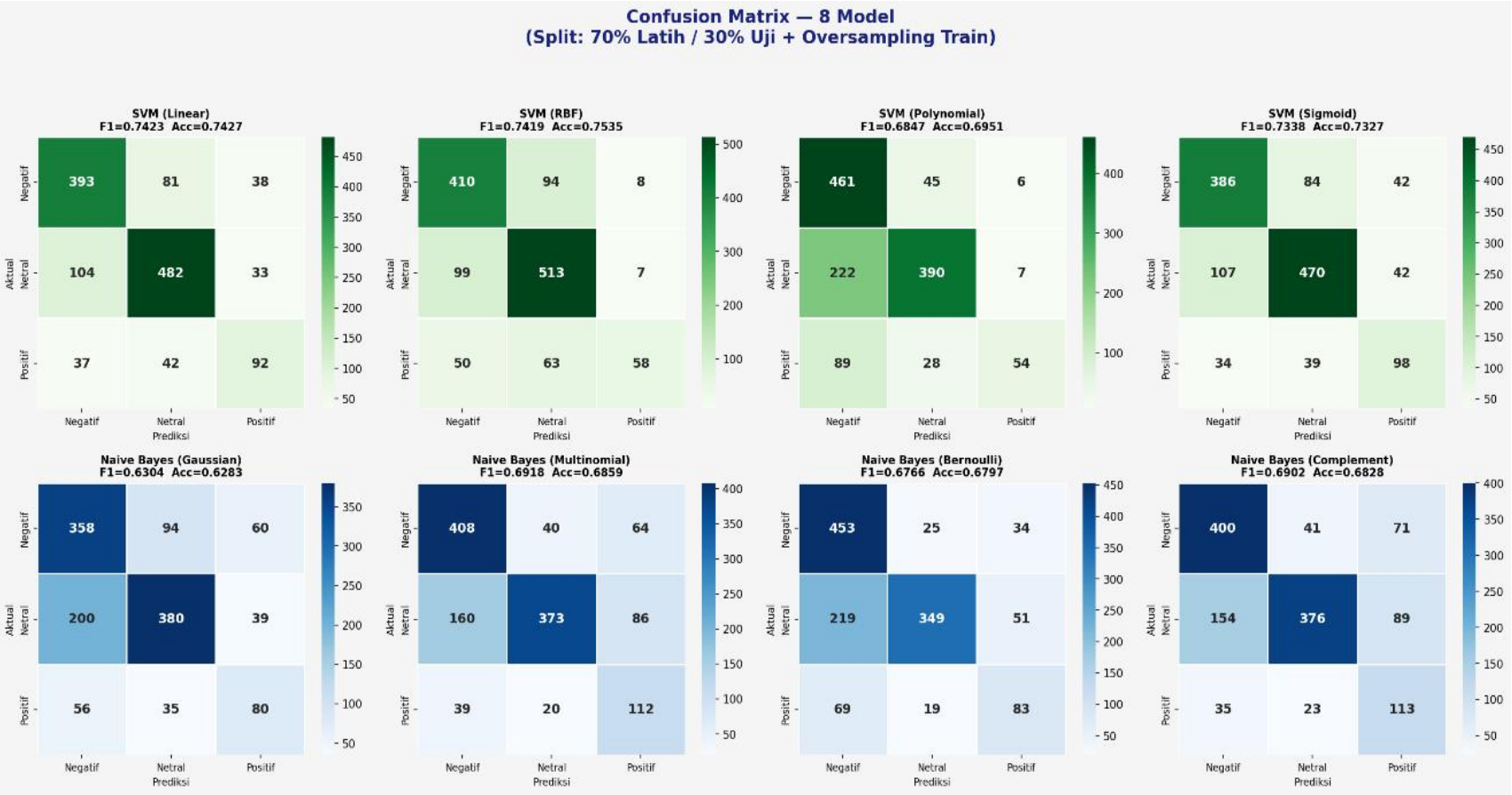
Tabel 3. Skenario 50% Latih dan 50% Uji

Kategori	Latih (Sebelum Oversample)	Latih (Sesudah Oversample)	Uji (Test)
Negatif	853	853	853
Netral	1031	1031	1031
Positif	285	285	286
Total	2169	2169	2170

Tabel 4. Skenario 30% Latih dan 70% Uji

Kategori	Latih (Sebelum Oversample)	Latih (Sesudah Oversample)	Uji (Test)
Negatif	512	512	1194
Netral	618	618	1444
Positif	171	171	400
Total	1301	1301	3038

Naïve Bayes dan SVM 70% Latih : 30 % Uji

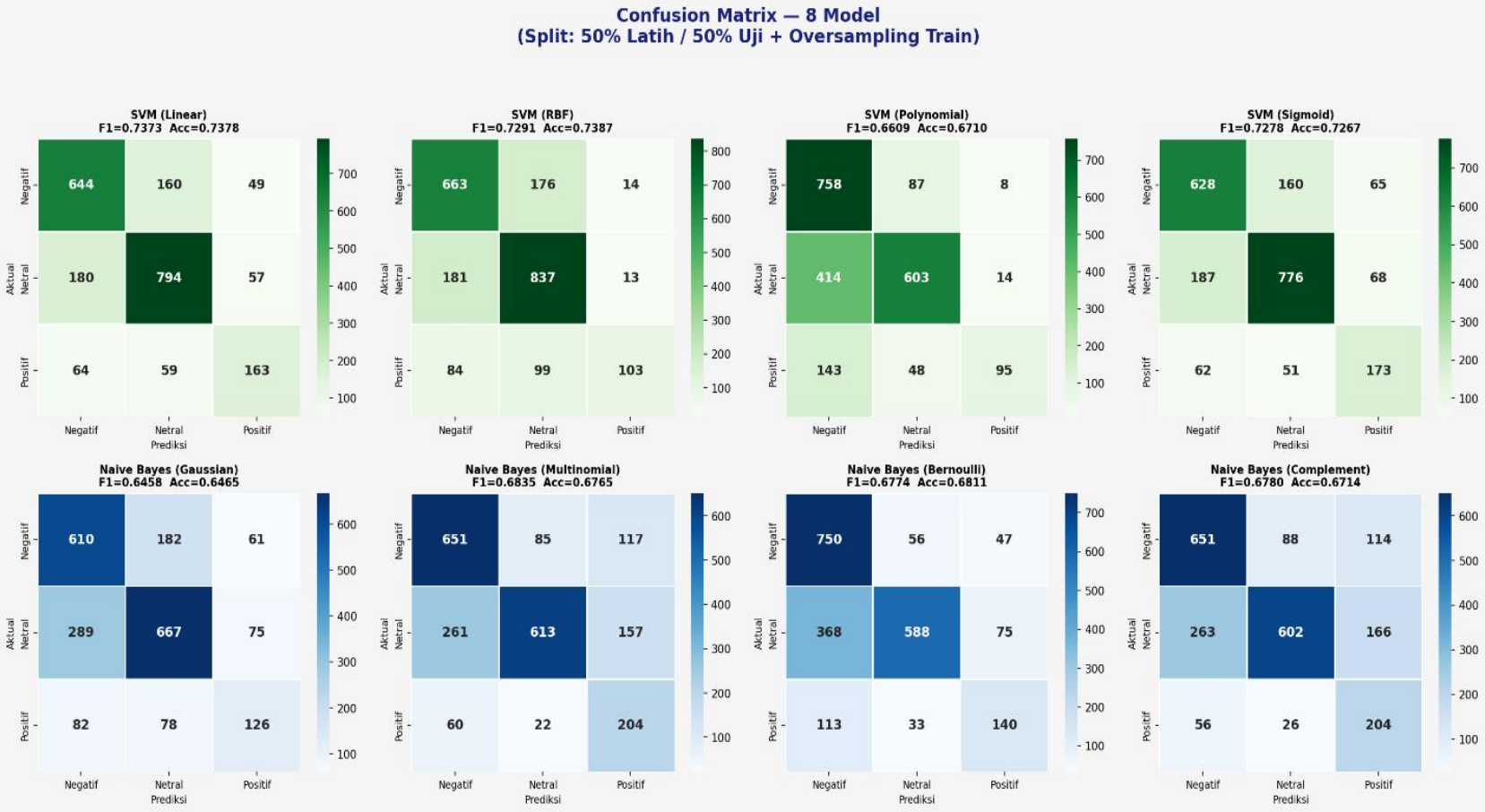


Pada skenario pembagian data 70% data latih dan 30% data uji, metode SVM memberikan performa yang lebih baik dibandingkan Naïve Bayes. SVM Linear memperoleh F1-Score tertinggi sebesar 74,23%, sedangkan SVM RBF mencapai Accuracy tertinggi sebesar 75,35%. Pada kelompok Naïve Bayes, Multinomial Naïve Bayes menjadi model terbaik dengan F1-Score 69,18% dan Accuracy 68,59%, sedangkan Gaussian Naïve Bayes menunjukkan performa terendah. Hasil confusion matrix juga menunjukkan bahwa SVM mampu mengklasifikasikan data sentimen lebih baik, terutama pada kelas netral.

Naïve Bayes dan SVM 70% Latih : 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%	-
Naïve Bayes	Complement	69,02%	68,28%	-
Naïve Bayes	Bernoulli	67,66%	67,97%	-
Naïve Bayes	Gaussian	63,04%	62,83%	Terendah

Naïve Bayes dan SVM 50% Latih : 50 % Uji

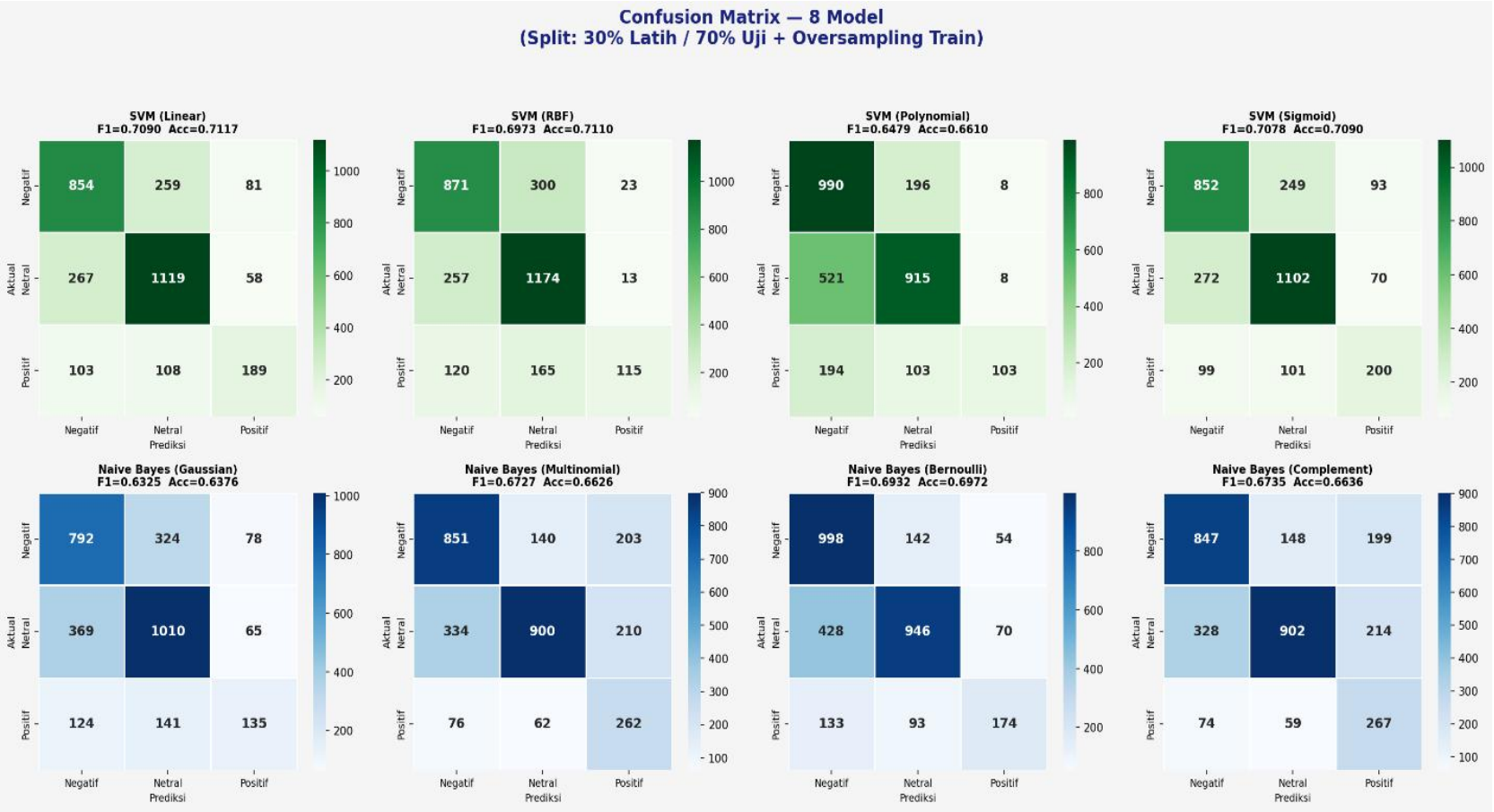


Pada skenario pembagian data 50% data latih dan 50% data uji, metode SVM kembali menunjukkan performa yang lebih baik dibandingkan Naïve Bayes. SVM Linear memperoleh F1-Score tertinggi sebesar 73,73%, sedangkan SVM RBF mencapai Accuracy tertinggi sebesar 73,87%. Pada kelompok Naïve Bayes, Multinomial Naïve Bayes memperoleh F1-Score terbaik sebesar 68,35%, sedangkan Bernoulli Naïve Bayes memperoleh Accuracy tertinggi sebesar 68,11%. Hasil confusion matrix juga menunjukkan bahwa SVM mampu mengklasifikasikan data sentimen dengan lebih baik, terutama pada kelas netral.

Naïve Bayes dan SVM 50% Latih : 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%	-
Naïve Bayes	Complement	67,80%	67,14%	-
Naïve Bayes	Bernoulli	67,74%	68,11%	-
Naïve Bayes	Gaussian	64,58%	64,65%	Terendah

Naïve Bayes dan SVM 30% Latih : 70 % Uji



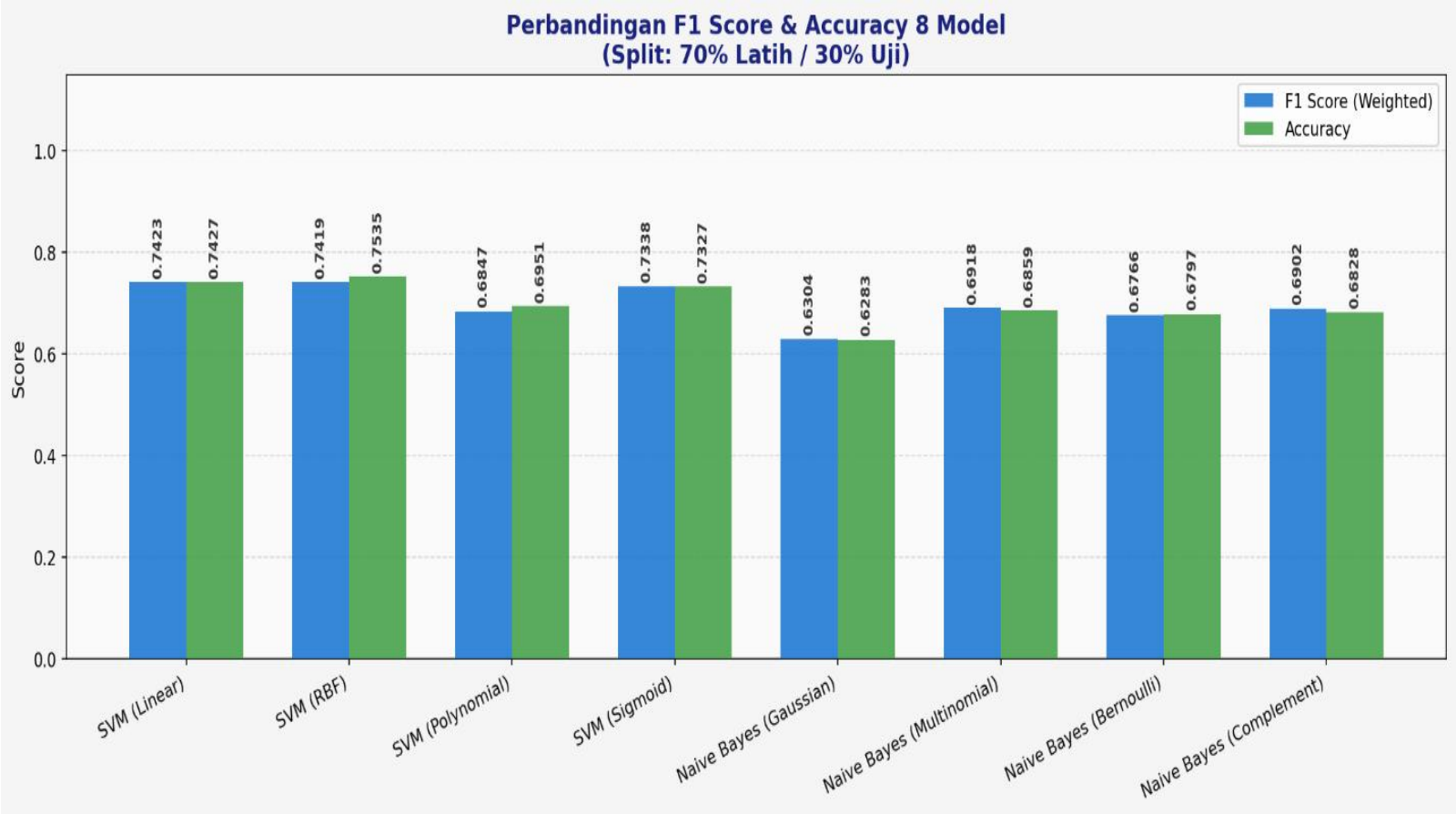
Pada skenario pembagian data 30% data latih dan 70% data uji, metode SVM tetap memberikan performa terbaik dibandingkan Naïve Bayes. SVM Linear memperoleh F1-Score tertinggi sebesar 70,90% sekaligus Accuracy tertinggi sebesar 71,17%. Pada kelompok Naïve Bayes, Bernoulli Naïve Bayes menjadi model terbaik dengan F1-Score sebesar 69,32% dan Accuracy sebesar 69,72%, sedangkan Gaussian Naïve Bayes menunjukkan performa terendah. Hasil confusion matrix juga menunjukkan bahwa SVM mampu mengklasifikasikan data sentimen dengan lebih baik dibandingkan Naïve Bayes.

Naïve Bayes dan SVM 30% Latih : 70 % Uji

Model	Kernel / Varian	F1-Score (Weighted)	Accuracy	Keterangan
SVM	Linear	70,90%	71,17%	Terbaik F1
SVM	RBF	69,73%	71,10%	Terbaik Acc
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%	-
Naïve Bayes	Gaussian	63,25%	63,76%	Terendah

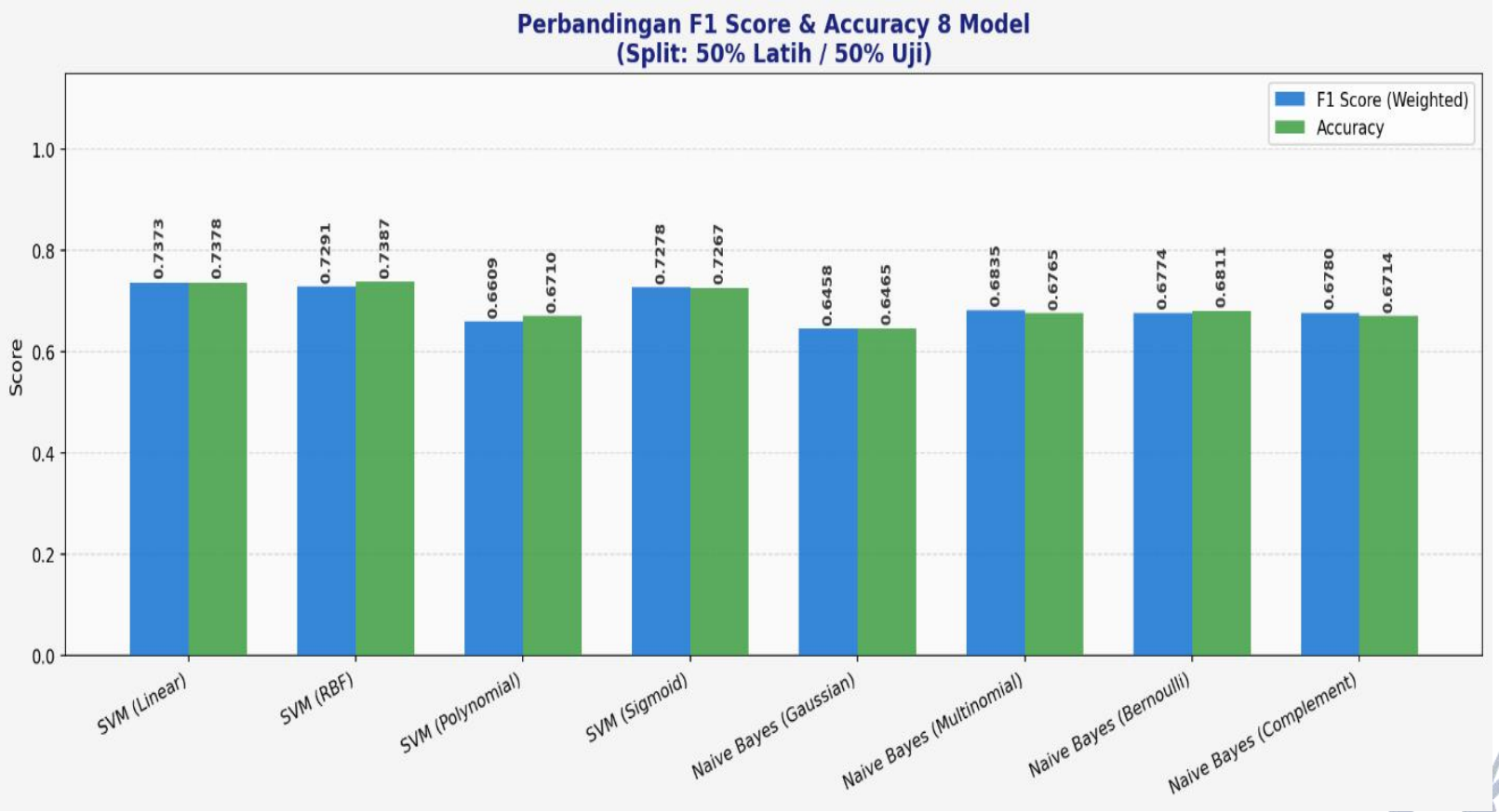
Perbandingan F1-Score dan Accuracy Naïve Bayes dan SVM 70% Latih : 30% Uji

Berdasarkan Gambar, model SVM menunjukkan performa yang lebih baik dibandingkan Naïve Bayes pada skenario pembagian data 70% data latih dan 30% data uji. SVM Linear memperoleh F1-Score tertinggi sebesar 74,23%, sedangkan Accuracy tertinggi dicapai oleh SVM RBF sebesar 75,35%. Pada kelompok Naïve Bayes, Multinomial Naïve Bayes memperoleh F1-Score tertinggi sebesar 69,18%, sedangkan Bernoulli Naïve Bayes mencapai Accuracy tertinggi sebesar 67,97%. Hasil ini menunjukkan bahwa SVM, khususnya kernel Linear dan RBF, memberikan performa klasifikasi sentimen yang lebih baik dibandingkan seluruh varian Naïve Bayes.



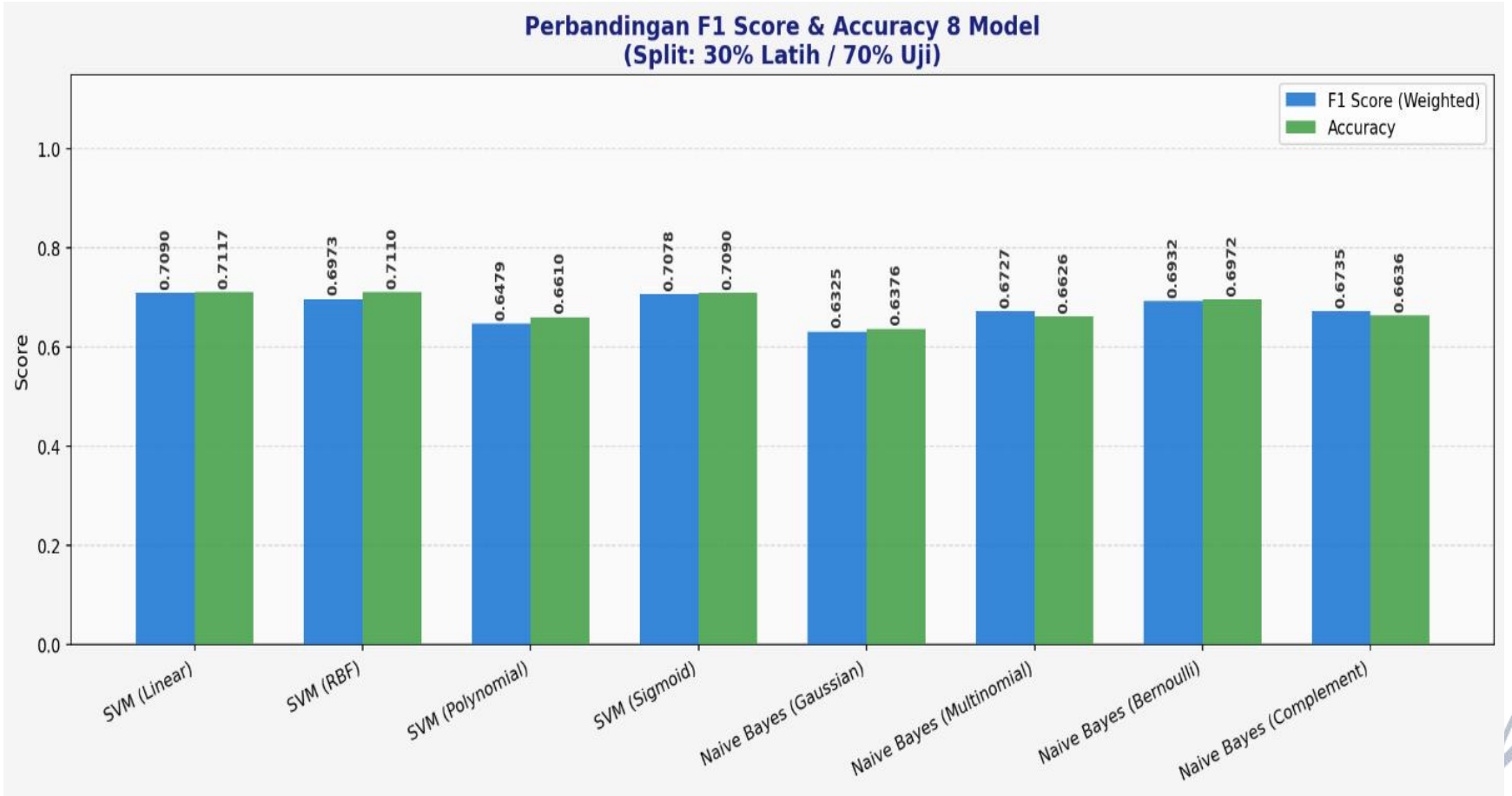
Perbandingan F1-Score dan Accuracy Naïve Bayes dan SVM 50% Latih : 50% Uji

Berdasarkan Gambar, model SVM tetap menunjukkan performa yang lebih baik dibandingkan Naïve Bayes pada skenario pembagian data 50% data latih dan 50% data uji. SVM Linear memperoleh F1-Score tertinggi sebesar 73,73%, sedangkan Accuracy tertinggi dicapai oleh SVM RBF sebesar 73,87%. Pada kelompok Naïve Bayes, Multinomial Naïve Bayes memperoleh F1-Score tertinggi sebesar 68,35%, sedangkan Bernoulli Naïve Bayes mencapai Accuracy tertinggi sebesar 68,11%. Meskipun terjadi sedikit penurunan dibandingkan skenario 70% data latih dan 30% data uji, SVM tetap memberikan performa klasifikasi yang lebih baik dibandingkan seluruh varian Naïve Bayes.



Perbandingan F1-Score dan Accuracy Naïve Bayes dan SVM 30% Latih : 70% Uji

Berdasarkan Gambar, model SVM tetap menunjukkan performa yang lebih baik dibandingkan Naïve Bayes pada skenario pembagian data 30% data latih dan 70% data uji. SVM Linear memperoleh F1-Score tertinggi sebesar 70,90% sekaligus Accuracy tertinggi sebesar 71,17%. Pada kelompok Naïve Bayes, Bernoulli Naïve Bayes menjadi model terbaik dengan F1-Score sebesar 69,32% dan Accuracy sebesar 69,72%. Meskipun performa seluruh model mengalami sedikit penurunan akibat berkurangnya jumlah data latih, SVM tetap memberikan hasil klasifikasi yang lebih baik dibandingkan seluruh varian Naïve Bayes.



Rekap 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%

Analisis Word Cloud

Kelas: Negatif



Kelas: Netral



Kelas: Positif



- **Sentimen Negatif**: Didominasi kata mati, air, daerah, rumah, layanan, dan distribusi. Kata mati paling sering muncul, menunjukkan banyak keluhan terkait gangguan distribusi, kualitas air, dan layanan PDAM. Kemunculan kata tolong dan info mencerminkan harapan masyarakat terhadap perbaikan layanan.
- **Sentimen Netral**: Didominasi kata air, bersih, mati, keruh, langgan, rumah, daerah, dan tagih. Tweet pada kelas ini umumnya berisi informasi mengenai kondisi layanan, distribusi air, penggunaan air, serta pembayaran tagihan tanpa menunjukkan opini yang kuat.
- **Sentimen Positif**: Didominasi kata alir, bersih, terima, layanan, masyarakat, dan distribusi. Kata terima, bantu, cepat, dan respon menunjukkan apresiasi masyarakat terhadap kualitas air, kelancaran distribusi, serta respons layanan PDAM yang dinilai baik.

Kesimpulan

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, 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 didominasi oleh 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.


UMSIDA
DARI SINI PENCERAHAN BERSEMI