

ANALISIS SENTIMEN TERHADAP MENTERI KEUANGAN PURBAYA YUDHI SADEWA PADA MEDIA SOSIAL X MENGGUNAKAN *NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINE*

Muhammad Dzakil Aqli^{1*}, Nina Sulistiyowati², Apriade Voutama³

^{1,2,3}Universitas Singaperbangsa Karawang; Jl.H.S. Ronggowaluyo, Kel. Puseurjaya, Kec. Telukjambe Timur, Kab. Karawang, Jawa Barat; 41361

Keywords:

Analisis Sentimen, *Naïve Bayes*, *Support Vector Machine*, Media Sosial X, KDD

Correspondent Email:

dzakilaqli@gmail.com

Abstrak. Perkembangan media sosial menjadikan platform X sebagai media utama masyarakat dalam menyampaikan opini terhadap tokoh publik, termasuk Menteri Keuangan Purbaya Yudhi Sadewa. Penelitian ini bertujuan menganalisis sentimen publik terhadap Purbaya Yudhi Sadewa serta membandingkan kinerja algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM). Metode yang digunakan adalah *Knowledge Discovery in Database* (KDD) yang meliputi *data selection*, *preprocessing*, *labeling*, *transformation*, *data mining*, dan *evaluation*. Data diperoleh melalui *scraping* sebanyak 3.173 komentar dan diseleksi menjadi 2.021 data relevan. Setelah pelabelan sentimen menggunakan *InSet Lexicon* dan penghapusan data netral, diperoleh 1.901 data untuk klasifikasi. Transformasi data dilakukan menggunakan TF-IDF dan penanganan ketidakseimbangan kelas menggunakan SMOTE. Hasil penelitian menunjukkan sentimen positif mendominasi sebesar 77,12% dan sentimen negatif sebesar 22,88%. Algoritma SVM menghasilkan performa terbaik dengan akurasi 90,29%, *precision* 90,02%, *recall* 90,29%, dan *F1-score* 90,05%, sedangkan *Naïve Bayes* memperoleh akurasi terbaik 83,99%. Hasil penelitian menunjukkan bahwa SVM lebih efektif dibandingkan *Naïve Bayes* dalam klasifikasi sentimen pada media sosial X.



Copyright © 2026 The Author(s), Published by Jurnal Informatika dan Teknik Elektro Terapan. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Abstract. The rapid growth of social media has made X a major platform for expressing public opinions regarding public figures, including the Minister of Finance, Purbaya Yudhi Sadewa. This study aims to analyze public sentiment toward Purbaya Yudhi Sadewa and compare the performance of the *Naïve Bayes* and *Support Vector Machine* (SVM) algorithms. The research employed the *Knowledge Discovery in Database* (KDD) methodology, consisting of *data selection*, *preprocessing*, *labeling*, *transformation*, *data mining*, and *evaluation*. A total of 3,173 comments were collected and filtered into 2,021 relevant comments. After sentiment labeling using the *InSet Lexicon* and removing neutral data, 1,901 comments were used for classification. TF-IDF was applied for feature extraction, while SMOTE was used to address class imbalance. The results show that positive sentiment dominated the dataset (77.12%), while negative sentiment accounted for 22.88%. SVM achieved the best performance with 90.29% accuracy, 90.02% precision, 90.29% recall, and 90.05% F1-score, whereas *Naïve Bayes* achieved its best accuracy of 83.99%. These findings indicate that SVM is more effective than *Naïve Bayes* for sentiment classification on social media X.

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong pertumbuhan media sosial sebagai sarana komunikasi dan pertukaran informasi yang digunakan secara luas oleh masyarakat [1]. Salah satu platform yang banyak digunakan untuk menyampaikan opini publik adalah media sosial X (sebelumnya Twitter), yang memungkinkan pengguna berbagi pandangan terhadap berbagai isu sosial, ekonomi, maupun politik secara cepat dan terbuka [2].

Media sosial X juga menjadi ruang diskusi yang aktif terkait tokoh publik dan kebijakan pemerintah. Salah satu tokoh yang banyak diperbincangkan adalah Menteri Keuangan Purbaya Yudhi Sadewa sejak menjabat pada September 2025. Berbagai kebijakan ekonomi yang diterapkannya menimbulkan beragam tanggapan dari masyarakat, baik berupa dukungan maupun kritik [3]. Besarnya jumlah opini yang dihasilkan menjadikan media sosial X sebagai sumber data yang relevan untuk mengukur persepsi publik secara objektif melalui analisis sentimen.

Analisis sentimen merupakan teknik text mining yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini pengguna ke dalam kategori positif, negatif, atau netral berdasarkan isi teks [4]. Dalam penelitian analisis sentimen, algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM) merupakan dua metode klasifikasi yang banyak digunakan karena memiliki performa yang baik dalam pengolahan data teks [5], [6]. Penelitian sebelumnya menunjukkan bahwa SVM cenderung menghasilkan akurasi yang lebih tinggi pada data berdimensi besar, sedangkan *Naïve Bayes* memiliki keunggulan dari sisi kesederhanaan dan efisiensi komputasi [7].

Meskipun berbagai penelitian telah membandingkan kedua algoritma tersebut, penelitian mengenai sentimen publik terhadap Menteri Keuangan Purbaya Yudhi Sadewa masih sangat terbatas. Selain itu, belum banyak penelitian yang membandingkan performa *Naïve Bayes* dan SVM pada data opini masyarakat terkait figur tersebut dengan penerapan TF-IDF dan SMOTE sebagai penanganan ketidakseimbangan kelas. Oleh karena itu, penelitian ini dilakukan untuk menganalisis sentimen publik terhadap Purbaya Yudhi Sadewa pada media sosial X serta membandingkan kinerja algoritma *Naïve Bayes*

dan *Support Vector Machine* dalam klasifikasi sentimen.

2. TINJAUAN PUSTAKA

2.1 Analisis Sentimen

Analisis sentimen merupakan metode pengolahan data berbasis teks yang memanfaatkan *Natural Language Processing* (NLP) dan *Machine Learning* (ML) untuk mengidentifikasi serta mengklasifikasikan opini ke dalam kategori positif, negatif, atau netral [8]. Pada media sosial, analisis sentimen banyak digunakan untuk memahami persepsi publik terhadap suatu isu, produk, layanan, maupun tokoh tertentu [9].

2.2 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan proses penemuan pengetahuan dari data yang meliputi tahapan *selection*, *preprocessing*, *transformation*, *data mining*, dan *evaluation*. Metode ini digunakan sebagai kerangka kerja penelitian untuk mengolah data komentar pada media sosial X hingga menghasilkan informasi sentimen publik yang dapat dianalisis lebih lanjut [10].

2.3 Term Frequency-Inverse Document Frequency (TF-IDF)

TF-IDF merupakan metode pembobotan kata yang digunakan untuk mengubah data teks menjadi representasi numerik berdasarkan tingkat kepentingan suatu kata dalam dokumen. Metode ini banyak digunakan pada penelitian *text mining* dan analisis sentimen karena mampu meningkatkan kualitas representasi fitur teks [11].

2.4 Naïve Bayes

Naïve Bayes merupakan algoritma klasifikasi berbasis probabilitas yang menggunakan *Teorema Bayes* dengan asumsi independensi antar fitur. Algoritma ini banyak diterapkan dalam analisis sentimen karena memiliki proses komputasi yang sederhana dan efisien [12].

2.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma *machine learning* yang bekerja dengan mencari *hyperplane* optimal untuk memisahkan data ke dalam kelas yang

berbeda. SVM dikenal memiliki kemampuan yang baik dalam menangani data berdimensi tinggi seperti data teks sehingga sering digunakan dalam penelitian analisis sentimen [13].

2.6 Synthetic Minority Over-sampling Technique (SMOTE)

SMOTE merupakan metode oversampling yang digunakan untuk mengatasi ketidakseimbangan kelas dengan menghasilkan data sintesis pada kelas minoritas. Teknik ini bertujuan meningkatkan kemampuan model dalam mengenali seluruh kelas secara lebih seimbang [14].

2.7 Purbaya Yudhi Sadewa

Purbaya Yudhi Sadewa merupakan ekonom Indonesia yang menjabat sebagai Menteri Keuangan Republik Indonesia sejak September 2025. Sebagai pejabat publik yang memiliki peran strategis dalam pengelolaan kebijakan fiskal dan ekonomi nasional, berbagai kebijakan serta pernyataannya sering menjadi perhatian masyarakat dan memunculkan beragam tanggapan di media sosial. Tingginya intensitas diskusi publik terkait Purbaya Yudhi Sadewa menjadikannya objek yang relevan untuk dianalisis menggunakan pendekatan analisis sentimen guna mengetahui kecenderungan persepsi masyarakat terhadap figur dan kebijakan yang diambilnya [15].

2.8 Lexicon-Based

Lexicon-Based merupakan metode analisis sentimen yang menentukan polaritas teks berdasarkan kamus sentimen yang berisi kata positif dan negatif beserta bobot nilainya. Metode ini digunakan untuk memberikan label sentimen secara otomatis dengan menghitung total skor dari kata-kata yang terdapat dalam dokumen [16].

Perhitungan skor sentimen dapat dirumuskan sebagai berikut:

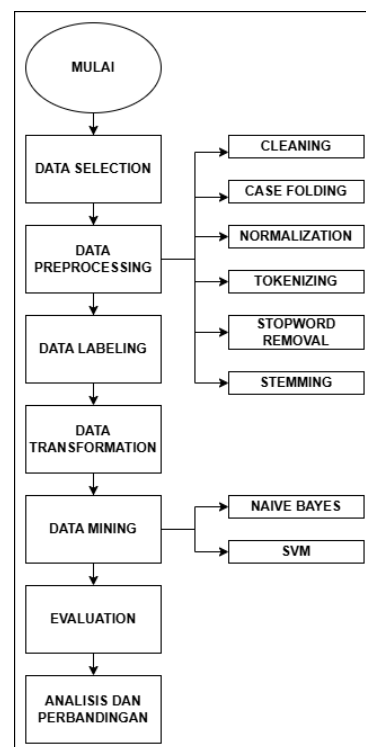
$$\text{Skor} = \sum (\text{skor positif}) - \sum (\text{skor negatif})$$

Perhitungan skor sentimen dilakukan dengan menjumlahkan seluruh skor kata positif dan mengurangi jumlah skor kata negatif yang terdapat dalam dokumen. Nilai skor yang

diperoleh digunakan untuk menentukan polaritas sentimen suatu dokumen. Jika skor > 0 maka sentimen positif, skor < 0 maka sentimen negatif, dan skor $= 0$ maka sentimen netral [16].

3. METODE PENELITIAN

Penelitian ini menggunakan metode *Knowledge Discovery in Database (KDD)* untuk menganalisis sentimen publik terhadap Menteri Keuangan Purbaya Yudhi Sadewa pada media sosial X. Metode KDD dipilih karena mampu mengolah data teks tidak terstruktur menjadi informasi yang dapat digunakan dalam proses klasifikasi sentimen. Tahapan penelitian terdiri atas *data selection*, *preprocessing*, *labeling*, *transformation*, *data mining*, *evaluation*, dan *analisis dan perbandingan*.



Gambar 3.1 Alur Penelitian

3.1. Data Selection

Pada tahap ini dilakukan pengumpulan data dari media sosial X menggunakan teknik *scraping* dengan kata kunci “Purbaya”. Data diambil pada periode 8 September 2025 hingga 31 Maret 2026. Alat bantu yang digunakan yaitu Google Colab dan *tools Tweet-Harvest*. Data yang telah diperoleh kemudian disimpan dalam format CSV

3.2. Data Preprocessing

Tahap *preprocessing* bertujuan membersihkan dan menyiapkan data sebelum proses analisis. Proses yang dilakukan meliputi *cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming*. Tahapan ini menghasilkan data teks yang lebih terstruktur dan siap digunakan pada proses klasifikasi.

3.3. Data Labeling

Pelabelan sentimen dilakukan menggunakan metode *Lexicon-Based* dengan memanfaatkan kamus sentimen bahasa Indonesia *InSet Lexicon*. Hasil pelabelan kemudian divalidasi oleh ahli bahasa untuk memastikan kesesuaian label sentimen yang diberikan. Setelah data netral dihapus, diperoleh 1.901 data yang digunakan pada tahap klasifikasi.

3.4. Data Transformation

Data teks yang telah diberi label ditransformasikan ke dalam bentuk numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata berdasarkan tingkat kepentingannya dalam dokumen.

3.5. Data Mining

Tahap *data mining* dilakukan dengan membangun model klasifikasi menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM). Untuk mengatasi ketidakseimbangan kelas, diterapkan metode *Synthetic Minority Over-sampling Technique* (SMOTE) pada data latih. Pengujian dilakukan menggunakan rasio data latih dan data uji sebesar 90:10, 80:20, 70:30, dan 60:40.

3.6. Evaluation

Evaluasi model dilakukan menggunakan *Confusion Matrix* dan *Classification Report*. Kinerja model diukur berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui kemampuan masing-masing algoritma dalam mengklasifikasikan sentimen.

3.7. Analisis dan Perbandingan

Tahap analisis dan perbandingan dilakukan untuk mengevaluasi hasil klasifikasi sentimen yang diperoleh dari algoritma *Naïve*

Bayes dan *Support Vector Machine* (SVM). Perbandingan dilakukan berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score* yang dihasilkan pada setiap skenario pengujian. Hasil perbandingan digunakan untuk menentukan algoritma yang memiliki performa terbaik dalam mengklasifikasikan sentimen publik terhadap Menteri Keuangan Purbaya Yudhi Sadewa pada media sosial X.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Data penelitian diperoleh melalui proses *scraping* pada media sosial X menggunakan kata kunci "Purbaya" selama periode 8 September 2025 hingga 31 Maret 2026. Proses pengambilan data dilakukan menggunakan Google Colab dan *Tweet-Harvest*. Dari hasil *scraping* diperoleh sebanyak 3.173 komentar. Setelah dilakukan seleksi atribut dan penghapusan data yang tidak relevan, diperoleh 2.021 data yang digunakan pada penelitian ini.

4.2. Data Preprocessing

Tahap *preprocessing* dilakukan untuk membersihkan data sebelum proses klasifikasi. Tahapan yang dilakukan meliputi *cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming*. Proses ini bertujuan menghilangkan karakter yang tidak diperlukan, menyeragamkan format teks, serta mengubah kata menjadi bentuk dasar sehingga dapat meningkatkan kualitas data yang digunakan pada tahap klasifikasi.

Tabel 4.1 Hasil *Cleaning*

Sebelum	Sesudah
Kebijakan aneh ga adil & ga ada prioritas... Mengelola uang negara dg ugal2an dimanaaa Purbaya yg koar2... mana pendukungnya????	Kebijakan aneh ga adil amp ga ada prioritas Mengelola uang negara dg ugalan dimanaaa Purbaya yg koar mana pendukungnya

Tabel 4.2 Hasil *Case Folding*

Sebelum	Sesudah
Kebijakan aneh ga adil amp ga ada prioritas Mengelola uang negara dg ugalan dimanaaa Purbaya yg koar mana pendukungnya	kebijakan aneh ga adil amp ga ada prioritas mengelola uang negara dg ugalan dimanaaa purbaya yg koar mana pendukungnya

Tabel 4.3 Hasil *Normalization*

Sebelum	Sesudah
kebijakan aneh ga adil amp ga ada prioritas mengelola uang negara dg ugalan dimanaaa purbaya yg koar mana pendukungnya	kebijakan aneh tidak adil amp tidak ada prioritas mengelola uang negara dengan ugalan dimanaaa purbaya yang koar mana pendukungnya

Tabel 4.4 Hasil *Tokenizing*

Sebelum	Sesudah
kebijakan aneh tidak adil amp tidak ada prioritas mengelola uang negara dengan ugalan dimanaaa purbaya yang koar mana pendukungnya	['kebijakan', 'aneh', 'tidak', 'adil', 'amp', 'tidak', 'ada', 'prioritas', 'mengelola', 'uang', 'negara', 'dengan', 'ugalan', 'dimanaaa', 'purbaya', 'yang', 'koar', 'mana', 'pendukungnya']

Tabel 4.5 Hasil *Stopword Removal*

Sebelum	Sesudah
['kebijakan', 'aneh', 'tidak', 'adil', 'amp', 'tidak', 'ada', 'prioritas', 'mengelola', 'uang', 'negara', 'dengan', 'ugalan', 'dimanaaa', 'purbaya', 'yang', 'koar', 'mana', 'pendukungnya']	['kebijakan', 'aneh', 'adil', 'prioritas', 'mengelola', 'uang', 'negara', 'ugalan', 'dimanaaa', 'purbaya', 'koar', 'pendukungnya']

Tabel 4.6 Hasil *Stemming*

Sebelum	Sesudah
['kebijakan', 'aneh', 'adil', 'prioritas', 'mengelola', 'uang', 'negara', 'ugalan', 'dimanaaa', 'purbaya', 'koar', 'pendukungnya']	['kebijakan', 'aneh', 'adil', 'prioritas', 'kelola', 'uang', 'negara', 'ugalan', 'dimanaaa', 'purbaya', 'koar', 'dukung']

4.3. Data Labeling

Pelabelan sentimen dilakukan menggunakan metode *Lexicon-Based* dengan memanfaatkan kamus sentimen bahasa Indonesia *InSet Lexicon*. Hasil pelabelan kemudian divalidasi oleh ahli bahasa Indonesia. Dari 2.021 data yang diperoleh, terdapat 1.466 sentimen positif, 435 sentimen negatif, dan 120 sentimen netral. Data netral kemudian dihapus sehingga diperoleh 1.901 data yang digunakan dalam proses klasifikasi.

Tabel 4.7 Perhitungan *Lexicon*

Stemmi ng	Positiv e_hits	Negativ e_hits	Skor	Sentim ent
purbaya sosok cipta sistem ekonomi tara adil	(tara,2), (adil,4)	(sistem, -4)	2.0	Positif

4.4. Data Transformation

Tahap transformasi data dilakukan menggunakan metode TF-IDF untuk mengubah data teks menjadi representasi numerik. Metode ini memberikan bobot pada setiap kata berdasarkan tingkat kemunculannya dalam dokumen sehingga dapat digunakan sebagai masukan bagi algoritma klasifikasi.

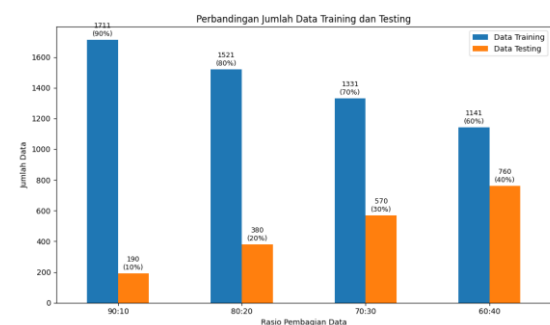
TF-IDF Matrix Shape: (1901, 1988)

	abal	abang	abdi	abdul	acara	aceh	acu	acung	ada	adab	adaptif	adem	adi	adil	
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.460865
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.433547
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.448600
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.478249
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.546867

Gambar 4.1 Hasil TF-IDF

4.5. Data Mining

Proses klasifikasi dilakukan menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM). Pengujian dilakukan menggunakan empat rasio pembagian data, yaitu 90:10, 80:20, 70:30, dan 60:40. Untuk mengatasi ketidakseimbangan kelas, diterapkan metode SMOTE pada data latih sehingga diperoleh model sebelum dan sesudah SMOTE.

**Gambar 4.2** Perbandingan Data Training dan Data Testing pada Setiap Rasio

4.6. Evaluation

Evaluasi model dilakukan menggunakan *Confusion Matrix* dan *Classification Report* dengan parameter *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa algoritma SVM menghasilkan performa yang lebih baik dibandingkan *Naïve Bayes* pada hampir seluruh skenario pengujian.

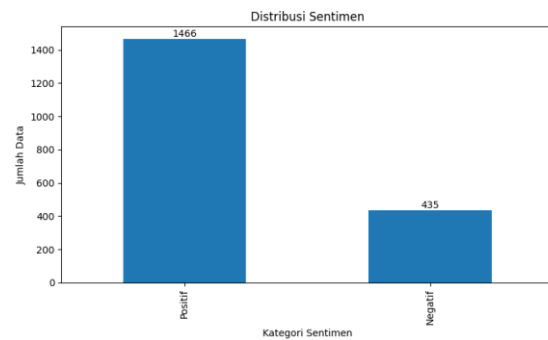
Rasio	Kondisi	Accuracy	Precision	Recall	F1-Score	Model
90:10	Sebelum SMOTE	76.44	59.14	76.44	66.69	Naive Bayes
90:10	Sesudah SMOTE	81.68	83.58	81.68	82.34	Naive Bayes
80:20	Sebelum SMOTE	77.95	78.40	77.95	69.48	Naive Bayes
80:20	Sesudah SMOTE	83.99	85.68	83.99	84.56	Naive Bayes
70:30	Sebelum SMOTE	78.11	80.16	78.11	69.77	Naive Bayes
70:30	Sesudah SMOTE	82.49	83.91	82.49	83.01	Naive Bayes
60:40	Sebelum SMOTE	77.79	79.55	77.79	68.93	Naive Bayes
60:40	Sesudah SMOTE	81.73	84.01	81.73	82.49	Naive Bayes
90:10	Sebelum SMOTE	88.48	88.17	88.48	87.80	SVM
90:10	Sesudah SMOTE	90.05	89.86	90.05	89.93	SVM
80:20	Sebelum SMOTE	90.29	90.02	90.29	90.05	SVM
80:20	Sesudah SMOTE	89.50	89.21	89.50	89.27	SVM
70:30	Sebelum SMOTE	89.67	89.35	89.67	89.30	SVM
70:30	Sesudah SMOTE	89.84	89.54	89.84	89.50	SVM
60:40	Sebelum SMOTE	88.70	88.31	88.70	88.18	SVM
60:40	Sesudah SMOTE	89.88	89.79	89.88	89.83	SVM

Gambar 4.3 Perbandingan Hasil *Naïve Bayes* dan SVM

Model terbaik SVM diperoleh pada rasio 80:20 sebelum SMOTE dengan akurasi 90,29%, precision 90,02%, recall 90,29%, dan F1-score 90,05%. Sementara itu, model terbaik *Naïve Bayes* diperoleh pada rasio 80:20 setelah SMOTE dengan akurasi 83,99%, precision 85,68%, recall 83,99%, dan F1-score 84,56%.

4.7. Analisis dan Perbandingan

Berdasarkan hasil pelabelan sentimen, opini publik terhadap Menteri Keuangan Purbaya Yudhi Sadewa didominasi oleh sentimen positif sebesar 77,12%, sedangkan sentimen negatif sebesar 22,88%. Dominasi sentimen positif menunjukkan bahwa sebagian besar pengguna media sosial X memberikan tanggapan yang cenderung mendukung terhadap sosok maupun kebijakan yang dikaitkan dengan Purbaya Yudhi Sadewa selama periode penelitian.



Gambar 4.4 Visualisasi Distribusi Sentimen

Hasil perbandingan menunjukkan bahwa algoritma SVM memiliki performa yang lebih unggul dibandingkan *Naïve Bayes*. Hal ini terlihat dari nilai akurasi, *precision*, *recall*, dan *F1-score* yang lebih tinggi. Kemampuan SVM dalam menangani data teks berdimensi tinggi hasil pembobotan TF-IDF menjadi salah satu faktor yang memengaruhi tingginya performa klasifikasi yang diperoleh.

Tabel 4.8 Perbandingan *Naïve Bayes* dan SVM Setiap Rasio

Metode	Teknik Penyeimbangan Data	Rasio	Confusion Matrix			
			Accuracy	Precision	Recall	F1-score
Naïve Bayes	Sebelum SMOTE	90:10	76.44	59.14	76.44	66.69
		80:20	77.95	78.40	77.95	69.48
		70:30	78.11	80.16	78.11	69.77
		60:40	77.79	79.55	77.79	68.93
		90:10	81.68	83.58	81.68	82.34
	Setelah SMOTE	80:20	83.99	85.68	83.99	84.56
		70:30	82.49	83.91	82.49	83.01
		60:40	81.73	84.01	81.73	82.49
		90:10	88.48	88.17	88.48	87.80
		80:20	90.29	90.02	90.29	90.05
Support Vector Machine (SVM)	Sebelum SMOTE	70:30	89.67	89.35	89.67	89.30
		60:40	88.70	88.31	88.70	88.18
		90:10	90.05	89.86	90.05	89.93
		80:20	89.50	89.21	89.50	89.27
		70:30	89.84	89.54	89.84	89.50
	Setelah SMOTE	60:40	89.88	89.79	89.88	89.83

5. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat ditarik beberapa kesimpulan sebagai berikut:

- Hasil analisis sentimen menunjukkan bahwa persepsi publik terhadap Menteri Keuangan Purbaya Yudhi Sadewa pada media sosial X cenderung positif. Dari 1.901 data yang digunakan dalam proses klasifikasi, sebanyak 1.466 data (77,12%) termasuk sentimen positif dan 435 data (22,88%) termasuk sentimen negatif.
- Proses klasifikasi sentimen berhasil dilakukan menggunakan tahapan *Knowledge Discovery in Database (KDD)*, meliputi *data selection*, *preprocessing*, pelabelan menggunakan

InSet Lexicon, transformasi data menggunakan TF-IDF, serta klasifikasi menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM).

- c. Berdasarkan hasil evaluasi, algoritma *Support Vector Machine* (SVM) menunjukkan performa terbaik dengan akurasi 90,29%, *precision* 90,02%, *recall* 90,29%, dan *F1-score* 90,05% pada rasio data 80:20 sebelum SMOTE. Sementara itu, model terbaik *Naïve Bayes* memperoleh akurasi 83,99%, *precision* 85,68%, *recall* 83,99%, dan *F1-score* 84,56% pada rasio 80:20 setelah SMOTE. Hasil tersebut menunjukkan bahwa SVM lebih efektif dibandingkan *Naïve Bayes* dalam mengklasifikasikan sentimen publik terhadap Menteri Keuangan Purbaya Yudhi Sadewa pada media sosial X.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah memberikan dukungan, bantuan, dan kontribusi selama pelaksanaan penelitian ini. Penghargaan juga disampaikan kepada semua pihak yang telah membantu dalam proses pengumpulan data, pengolahan data, serta penyusunan penelitian, sehingga penelitian ini dapat diselesaikan dengan baik dan sesuai dengan tujuan yang telah ditetapkan.

DAFTAR PUSTAKA

- [1] R. A. Putri, "Perkembangan Teknologi Informasi dan Dampaknya Terhadap Kehidupan Masyarakat," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 12, no. 2, pp. 45–52, 2023.
- [2] A. Kalalo, M. Parera, and J. Pieter, "Pemanfaatan Media Sosial sebagai Sarana Komunikasi Digital di Era Modern," *Jurnal Informatika dan Sistem Informasi*, vol. 8, no. 1, pp. 34–42, 2022.
- [3] M. Al Khaidar, "Purbaya Yudhi Sadewa dan Dinamika Kebijakan Fiskal Indonesia," *Jurnal Ekonomi dan Kebijakan Publik*, vol. 17, no. 2, pp. 101–110, 2025.
- [4] R. Tarigan and A. Yusupa, "Analisis Sentimen pada Media Sosial Menggunakan Pendekatan Machine Learning," *Jurnal Ilmu Komputer dan Informasi*, vol. 11, no. 3, pp. 210–218, 2024.
- [5] D. D. Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, pp. 73–79, 2022.
- [6] A. F. Rahman, A. Nurhadi, and S. Kurniawan, "Perbandingan Algoritma Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *Jurnal RESTI*, vol. 8, no. 1, pp. 45–53, 2024.
- [7] M. F. Pratama and E. Hidayat, "Evaluasi Kinerja Support Vector Machine pada Klasifikasi Teks Bahasa Indonesia," *Jurnal Teknologi dan Sistem Komputer*, vol. 12, no. 2, pp. 120–128, 2024.
- [8] T. N. Prakash and D. A. Aloysius, "Textual Sentiment Analysis Using Lexicon Based Approaches," *Annals of R.S.C.B.*, vol. 25, no. 4, pp. 13–18, 2021.
- [9] D. Purnamasari et al., *Pengantar Metode Analisis Sentimen*. 2024.
- [10] P. C. Behera, C. Dash, and S. Mohapatra, "Data Mining and Knowledge Discovery (KDD)," *IJRAR*, vol. 6, no. 1, pp. 35–47, 2019.
- [11] Y. Findawati and M. A. Rosid, *Buku Ajar Text Mining*. Sidoarjo: Umsida Press, 2020.
- [12] I. G. S. D. Putra and I. N. T. A. Putra, "Implementasi Metode Naïve Bayes Pada Analisis Sentimen Pengguna Aplikasi Mobile," vol. 13, no. 2, 2025.
- [13] M. Rafli Tanjung, N. Heryana, and S. Siska, "Analisis Sentimen Pengguna Platform X Terhadap Indonesia Emas 2045 Menggunakan Algoritma Support Vector Machine," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, pp. 7657–7665, 2024.
- [14] H. Permana, "Analisis Sentimen Terhadap Bakal Calon Presiden 2024 Dengan Algoritma Multinomial Naïve Bayes dan Oversampling SMOTE," vol. 7, no. 5, pp. 3257–3264, 2024.
- [15] R. L. Hakim, *Critical Reflection on Public Arguments: Purbaya Yudhi Sadewa*. 2025.
- [16] S. A. Nugraha, "Penerapan Lexicon Based Untuk Analisis Sentimen Masyarakat Indonesia Terhadap Danantara," vol. 9, no. 3, pp. 4949–4957, 2025.