

Analisis Sentimen Terhadap Kebijakan Purbaya Yudhi Sadewa Menggunakan *Support Vector Machine*

^{1*} Wildan Rayyana Muntaza, ²Patmi Kasih, ³Muh Aris Saputra

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ^{*1} wildanjr120hyper@gmail.com, ² patmikasih@unpkdr.ac.id, ³ aris.saputra@unpkdr.ac.id

Penulis Korespondensi : Wildan Rayyana Muntaza

Abstrak— *YouTube* merupakan salah satu media sosial yang banyak digunakan masyarakat untuk menyampaikan opini terhadap isu publik, termasuk kebijakan pemerintah. Banyaknya komentar yang dihasilkan menyebabkan proses analisis opini secara manual menjadi kurang efektif. Penelitian ini bertujuan untuk menganalisis sentimen komentar *YouTube* mengenai kebijakan Menteri Keuangan Purbaya Yudhi Sadewa menggunakan algoritma *Support Vector Machine* (SVM). Data penelitian diperoleh melalui proses *scraping* komentar dari beberapa kanal berita *YouTube*. Tahapan penelitian meliputi *preprocessing* data, pembobotan fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), klasifikasi menggunakan SVM, serta evaluasi menggunakan *Confusion Matrix*. Hasil pengujian menunjukkan bahwa model memperoleh akurasi sebesar 69,33%. Distribusi sentimen yang dihasilkan terdiri dari 56,3% sentimen positif, 31,1% sentimen negatif, dan 12,5% sentimen netral. Hasil penelitian menunjukkan bahwa mayoritas masyarakat memberikan tanggapan positif terhadap kebijakan Menteri Keuangan Purbaya Yudhi Sadewa dan algoritma SVM mampu digunakan untuk mengklasifikasikan sentimen komentar *YouTube* secara cukup baik.

Kata Kunci— Analisis Sentimen, SVM, TF-IDF, *YouTube*, Menteri Keuangan

Abstract— *YouTube* has become one of the most widely used social media platforms for expressing public opinions on government policies. The large number of comments makes manual opinion analysis inefficient. This study aims to analyze *YouTube* comments regarding the policies of the Indonesian Minister of Finance, Purbaya Yudhi Sadewa, using the *Support Vector Machine* (SVM) algorithm. The dataset was collected through comment *scraping* from several news channels on *YouTube*. The research stages included data *preprocessing*, feature weighting using *Term Frequency-Inverse Document Frequency* (TF-IDF), sentiment classification using SVM, and evaluation using a *Confusion Matrix*. The experimental results showed that the model achieved an accuracy of 69.33%. The sentiment distribution consisted of 56.3% positive, 31.1% negative, and 12.5% neutral sentiments. These results indicate that most public responses toward the Minister of Finance's policies were positive and that the SVM algorithm was able to perform sentiment classification on *YouTube* comments reasonably well.

Keywords— Sentiment Analysis, SVM, TF-IDF, *YouTube*, Minister of Finance

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Media sosial telah menjadi bagian penting dalam kehidupan masyarakat modern, khususnya sebagai sarana berbagi informasi dan menyampaikan opini secara *real-time* [1]. Platform *YouTube* menjadi salah satu media sosial yang banyak digunakan masyarakat untuk memberikan

tanggapan terhadap isu publik, termasuk kebijakan pemerintah. Komentar yang diberikan pengguna *YouTube* mengandung berbagai opini positif, negatif, maupun netral yang dapat dimanfaatkan sebagai sumber informasi untuk mengetahui persepsi publik [2].

Pergantian Menteri Keuangan Indonesia kepada Purbaya Yudhi Sadewa menimbulkan berbagai respon dari masyarakat di media sosial. Sebagai pejabat yang memiliki peran penting dalam kebijakan ekonomi dan fiskal negara, setiap kebijakan dan pernyataan yang disampaikan Menteri Keuangan menjadi perhatian publik. Banyak masyarakat menyampaikan pendapatnya melalui komentar pada kanal berita *YouTube* seperti Kompas TV, CNBC Indonesia, Metro TV, dan TvOne News. Banyaknya komentar tersebut menyebabkan proses analisis opini secara manual menjadi kurang efektif.

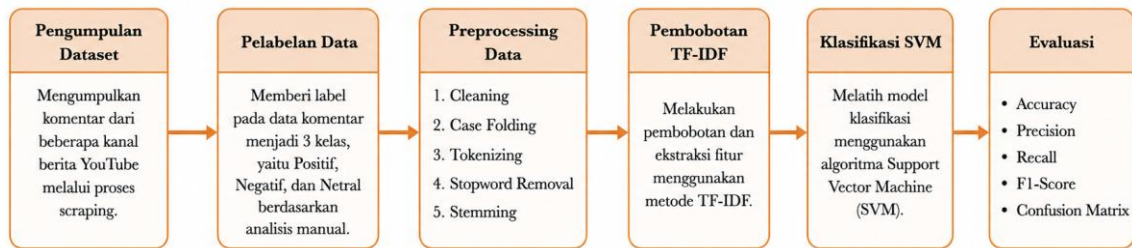
Pemanfaatan *machine learning* dalam analisis sentimen sangat diperlukan untuk membantu proses klasifikasi opini masyarakat secara otomatis. Analisis sentimen merupakan salah satu teknik *text mining* yang digunakan untuk mengelompokkan opini ke dalam kategori positif, negatif, dan netral [3]. Salah satu algoritma yang sering digunakan dalam klasifikasi teks adalah *Support Vector Machine* (SVM) karena memiliki performa yang baik pada data berdimensi tinggi dan menghasilkan tingkat akurasi yang tinggi [4].

Beberapa penelitian sebelumnya menunjukkan bahwa algoritma SVM memiliki performa yang baik dalam analisis sentimen media sosial. Penelitian oleh Muallafah dkk. berhasil memperoleh akurasi sebesar 95,02% pada analisis sentimen komentar *YouTube* menggunakan SVM [5]. Penelitian lain oleh Firdaus dkk. juga menunjukkan bahwa metode SVM mampu mencapai akurasi sebesar 96% dalam klasifikasi komentar berita *YouTube* [6]. Berdasarkan hasil penelitian tersebut, algoritma SVM dinilai cocok digunakan dalam penelitian ini.

Pada penelitian ini dilakukan analisis sentimen komentar *YouTube* mengenai kebijakan Menteri Keuangan Purbaya Yudhi Sadewa menggunakan algoritma *Support Vector Machine* (SVM). Penelitian ini bertujuan untuk mengetahui kecenderungan persepsi masyarakat terhadap kebijakan Menteri Keuangan berdasarkan komentar yang terdapat pada platform *YouTube*.

II. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan teknik analisis sentimen terhadap komentar *YouTube* mengenai kebijakan Menteri Keuangan Purbaya Yudhi Sadewa. Pendekatan kuantitatif digunakan karena penelitian berfokus pada proses pengolahan data teks, pengukuran hasil klasifikasi, serta evaluasi performa model menggunakan perhitungan statistik [7]. Proses penelitian dilakukan secara bertahap mulai dari pengumpulan data, *preprocessing*, pembobotan fitur, klasifikasi, hingga evaluasi model. Alur penelitian dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alir Analisis Sentimen

2.1 Dataset

Data yang digunakan pada penelitian ini berupa komentar pengguna *YouTube* yang berkaitan dengan kebijakan Menteri Keuangan Purbaya Yudhi Sadewa. Data diambil dari beberapa kanal berita *YouTube* seperti Kompas TV, CNBC Indonesia, Metro TV, TvOne News, dan Liputan6 menggunakan teknik *scraping*. Data ini di ambil dengan rentang waktu 14 september 2025 sampai dengan 18 mei 2026, dengan jumlah data 8.474. *Dataset* yang diperoleh kemudian disimpan dalam format CSV untuk memudahkan proses pengolahan data.

2.2 Pelabelan Data

Tahap pelabelan dilakukan untuk memberikan kategori sentimen pada komentar. Label sentimen yang digunakan terdiri dari positif, negatif, dan netral. Proses pelabelan dilakukan secara manual berdasarkan isi komentar agar hasil klasifikasi lebih sesuai dengan konteks opini masyarakat terhadap kebijakan Menteri Keuangan.

2.3 Preprocessing

Preprocessing merupakan tahap pembersihan dan transformasi data teks agar data yang awalnya tidak terstruktur dapat diproses oleh sistem *machine learning*. Pada penelitian ini, *preprocessing* dilakukan melalui beberapa tahapan, yaitu *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Tahap *cleaning* digunakan untuk menghapus URL, angka, tanda baca, emoji, dan karakter khusus pada komentar *YouTube*. Selanjutnya dilakukan *case folding* dengan mengubah seluruh huruf menjadi huruf kecil (*lowercase*). Setelah itu, proses *tokenizing* dilakukan untuk memecah kalimat menjadi kata-kata tunggal. Tahapan berikutnya adalah *stopword removal* untuk menghapus kata-kata umum yang tidak memiliki makna penting menggunakan *library* Sastrawi. Tahap terakhir yaitu *stemming* yang bertujuan mengubah kata berimbuhan menjadi kata dasar sehingga proses klasifikasi menjadi lebih optimal [8].

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

Setelah *preprocessing* dan pelabelan selesai dilakukan, data teks diubah menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)*. Metode *TF-IDF* digunakan untuk memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam dokumen dan keseluruhan *dataset* [9].

Perhitungan *TF-IDF* dapat dirumuskan pada persamaan (1).

$$IDF_t = \log \left(\frac{N}{df_t} \right) \quad (1)$$

Selanjutnya bobot *TF-IDF* dihitung menggunakan persamaan (2).

$$TF - IDF = tf \times \log \left(\frac{N}{df_t} \right) \quad (2)$$

2.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma *machine learning* yang digunakan untuk proses klasifikasi sentimen. SVM bekerja dengan mencari *hyperplane* terbaik untuk memisahkan data berdasarkan kelas sentimen [10]. Algoritma ini dipilih karena memiliki performa yang baik pada klasifikasi teks dan mampu menangani data berdimensi tinggi.

Fungsi keputusan pada SVM linear dapat dituliskan pada persamaan (3).

$$h(x) = w \cdot x + b \quad (3)$$

2.6 Evaluasi

Tahap evaluasi dilakukan untuk mengetahui performa model klasifikasi yang telah dibangun. Evaluasi menggunakan *Confusion Matrix* untuk menghitung nilai *accuracy*(4), *precision*(5), *recall*(6), dan *F1-score*(7) [11].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

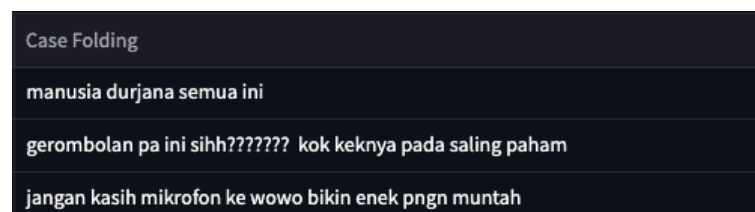
$$F1 \text{ Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

Hasil evaluasi digunakan untuk mengetahui tingkat keberhasilan algoritma SVM dalam mengklasifikasikan sentimen komentar *YouTube* terhadap kebijakan Menteri Keuangan Purbaya Yudhi Sadewa.

III. HASIL DAN PEMBAHASAN

3.1 Preprocessing

Tahap *preprocessing* dilakukan untuk mengubah data teks mentah menjadi data yang lebih terstruktur sebelum digunakan dalam proses klasifikasi. Tahapan yang dilakukan meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Hasil *preprocessing* menunjukkan adanya pengurangan rata-rata jumlah kata dari 16,86 kata menjadi 10,00 kata pada setiap komentar. Pengurangan jumlah kata tersebut menunjukkan bahwa proses *preprocessing* berhasil menghilangkan kata-kata yang tidak memiliki pengaruh signifikan terhadap proses klasifikasi. Contoh hasil transformasi *preprocessing* dapat dilihat pada Gambar 2, 3, 4, 5, dan 6.



Gambar 2. contoh hasil *case folding*

Cleaning
manusia durjana semua ini
gerombolan pa ini sihh kok keknya pada saling paham
jangan kasih mikrofon ke wowo bikin enek pngn muntah

Gambar 3. contoh hasil *cleaning*

Tokenizing
manusia, durjana, semua, ini
gerombolan, pa, ini, sihh, kok, keknya, pada, saling, paham
jangan, kasih, mikrofon, ke, wowo, bikin, enek, pngn, muntah

Gambar 4. contoh hasil *tokenizing*

Stopword Removal
manusia, durjana
gerombolan, pa, sihh, keknya, paham
kasih, mikrofon, wowo, bikin, enek, pngn, muntah

Gambar 5. contoh hasil *stopword removal*

Stemming
manusia durjana
gerombol pa sihh kek paham
kasih mikrofon wowo bikin enek pngn muntah

Gambar 2. contoh hasil *steming*

3.2 Klasifikasi Menggunakan SVM

Proses klasifikasi dilakukan menggunakan algoritma *Support Vector Machine (SVM)* dengan representasi fitur menggunakan *TF-IDF*. Model dilatih menggunakan data training dan kemudian diuji menggunakan data testing dengan perbandingan 90:10 untuk mengetahui kemampuan model dalam mengklasifikasikan sentimen komentar ke dalam kategori positif, negatif, dan netral. Hasil pengujian menunjukkan bahwa model memperoleh nilai akurasi sebesar 69,33%

3.3 Evaluasi Model

Evaluasi dilakukan menggunakan *Confusion Matrix* dan *Classification Report*. Hasil *Confusion Matrix* ditunjukkan pada Tabel 1.

Tabel 1. *Confusion Matrix*

Aktual/Prediksi	Negatif	Netral	Positif
Negatif	293	31	146
Netral	68	79	88
Positif	106	35	562

Berdasarkan hasil *Confusion Matrix* tersebut dapat diketahui bahwa model mampu mengklasifikasikan sentimen positif dengan cukup baik dibandingkan kategori lainnya.

Selanjutnya dilakukan pengukuran menggunakan metrik *precision*, *recall*, dan *f1-score* sebagaimana ditunjukkan pada Tabel 2.

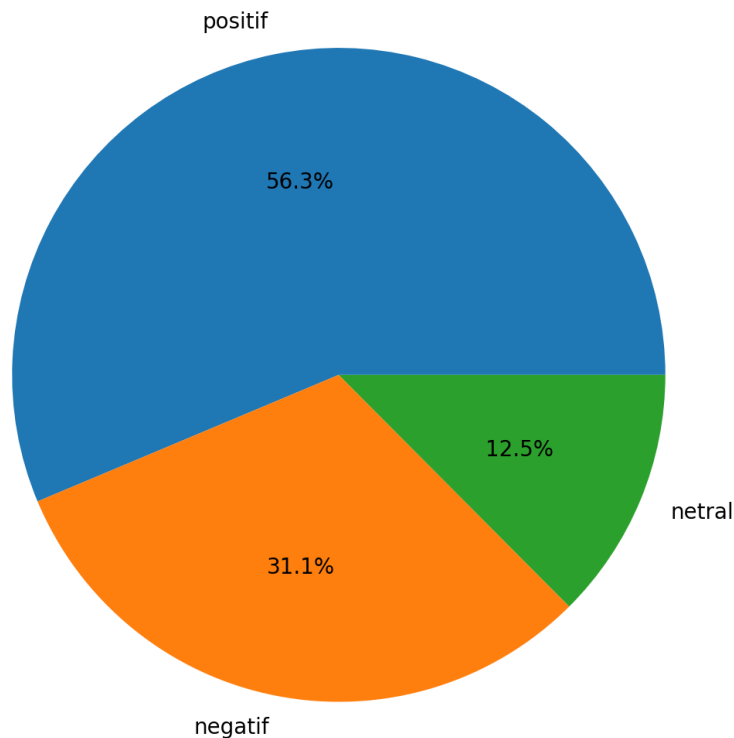
Tabel 2. *Classification Report*

Aktual/Prediksi	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Negatif	0.659	0.6442	0.6515
Netral	0.6	0.4565	0.5185
Positif	0.7331	0.7991	0.7646

Hasil evaluasi menunjukkan bahwa kategori positif memiliki performa terbaik dengan nilai *precision* sebesar 73,31%, *recall* sebesar 79,91%, dan *f1-score* sebesar 76,46%. Sementara itu kategori netral memiliki performa paling rendah karena karakteristik komentar netral cenderung memiliki kemiripan dengan komentar positif maupun negatif sehingga lebih sulit dibedakan oleh model.

3.4 Visualisasi Sentimen

Visualisasi hasil klasifikasi menunjukkan bahwa sentimen masyarakat terhadap kebijakan Menteri Keuangan Purbaya Yudhi Sadewa didominasi oleh sentimen positif. Dari keseluruhan data yang berhasil diklasifikasikan, diperoleh 56,3% sentimen positif, 31,1% sentimen negatif, dan 12,5% sentimen netral.



Gambar 3. Distribusi Sentimen Hasil Klasifikasi

Hasil tersebut menunjukkan bahwa sebagian besar masyarakat memberikan tanggapan yang cenderung mendukung kebijakan yang disampaikan oleh Menteri Keuangan, meskipun masih terdapat sejumlah komentar yang menunjukkan ketidaksetujuan maupun kritik terhadap kebijakan tersebut.

3.5 Analisis Wordcloud

Untuk mengetahui kata-kata yang sering muncul pada masing-masing kategori sentimen, dilakukan visualisasi menggunakan *wordcloud*. Hasil visualisasi menunjukkan bahwa pada sentimen negatif banyak muncul kata seperti rakyat, uang, dan pajak. Pada sentimen netral ditemukan kata yang sering muncul seperti orang, uang, dan jadi. Sementara pada sentimen positif didominasi kata jadi, rakyat, dan Indonesia.



Gambar 8. Wordcloud Sentimen Negatif, Netral, dan Positif

Visualisasi *wordcloud* memberikan gambaran mengenai topik pembahasan yang dominan pada setiap kategori sentimen sehingga dapat digunakan untuk memahami persepsi masyarakat secara lebih mendalam terhadap kebijakan Menteri Keuangan Purbaya Yudhi Sadewa.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *Support Vector Machine (SVM)* mampu digunakan untuk melakukan klasifikasi sentimen komentar *YouTube* terkait kebijakan Menteri Keuangan Purbaya Yudhi Sadewa. Proses analisis dilakukan melalui tahapan *preprocessing*, pembobotan fitur menggunakan *TF-IDF*, serta klasifikasi sentimen ke dalam tiga kategori, yaitu positif, negatif, dan netral.

Hasil pengujian menunjukkan bahwa model *SVM* memperoleh nilai akurasi sebesar 69,33% dengan nilai *precision*, *recall*, dan *F1-score* terbaik pada kategori sentimen positif. Berdasarkan hasil klasifikasi, sentimen masyarakat didominasi oleh sentimen positif sebesar 56,3%, diikuti sentimen negatif sebesar 31,1%, dan sentimen netral sebesar 12,5%. Hasil tersebut menunjukkan bahwa sebagian besar masyarakat memberikan tanggapan yang cenderung mendukung terhadap kebijakan Menteri Keuangan Purbaya Yudhi Sadewa.

Selain itu, visualisasi *wordcloud* menunjukkan bahwa kata-kata yang sering muncul pada setiap kategori sentimen memiliki karakteristik yang berbeda, sehingga dapat memberikan gambaran mengenai topik pembahasan yang dominan dalam persepsi masyarakat. Dengan demikian, analisis sentimen menggunakan algoritma *SVM* dapat dimanfaatkan sebagai salah satu metode untuk memahami opini publik terhadap kebijakan pemerintah secara lebih cepat dan objektif berdasarkan data yang diperoleh dari media sosial *YouTube*.

DAFTAR PUSTAKA

- [1] V. Agustina and A. Herliana, “Analisis Sentimen Publik atas Kebijakan Efisiensi Anggaran 2025 dengan Text Mining dan Natural Language Processing,” JURNAL MEDIA INFORMATIKA (JUMIN), vol. 6, no. 3, pp. 2182–2194, 2025.
- [2] R. Asrianto and M. Herwinanda, “Analisis Sentimen Kenaikan Harga Kebutuhan Pokok di Media Sosial Youtube Menggunakan Algoritma Support Vector Machine,” CoSciTech: Jurnal Computer Science and Information Technology, vol. 3, no. 3, pp. 431–440, 2022.
- [3] R. B. Dahlian and D. Sitanggang, “Analisis Sentimen Migrasi Televisi Digital pada Twitter Menggunakan Perbandingan Algoritma Multinomial Naïve Bayes, Support Vector Machines, dan Logistic Regression,” Jurnal SISFOKOM (Sistem Informasi dan Komputer), vol. 12, no. 2, pp. 280–288, 2023.
- [4] R. Firdaus, R. Al Hariri, and H. F. Amran, “Sentimen Analisis Masyarakat Tentang Penetapan Hari Raya Idul Adha Tahun 2023 Pada Video Youtube Menggunakan Algoritma Random Forest dan Support Vector Machine,” JURNAL FASILKOM, vol. 14, no. 1, pp. 278–285, 2024.
- [5] J. W. Iskandar and Y. Nataliani, “Perbandingan Naïve Bayes, SVM, dan k-NN untuk Analisis Sentimen Gadget Berbasis Aspek,” JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 5, no. 6, pp. 1120–1126, 2021.
- [6] J. Juliantono and Parjito, “Persepsi Publik Terhadap Kepemimpinan Firli Bahuri di KPK: Pendekatan Sentimen Twitter Dengan Naïve Bayes dan SVM,” JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika), vol. 10, no. 2, pp. 1272–1285, 2025.
- [7] D. Mualfah, Ramadhoni, R. Gunawan, and D. M. Suratno, “Analisis Sentimen Komentar YouTube TvOne Tentang Ustadz Abdul Somad Dideportasi Dari Singapura Menggunakan Algoritma SVM,” 2023.
- [8] S. Riyadi, L. K. Salsabila, C. Damarjati, and R. A. Karim, “Sentiment Analysis of YouTube Users on Blackpink Kpop Group Using IndoBERT,” 2024.
- [9] C. A. A. Soemedhy et al., “Analisis Komparasi Algoritma Machine Learning untuk Sentiment Analysis (Studi Kasus: Komentar YouTube Kekerasan Seksual),” 2022.
- [10] A. N. Syafia, M. F. Hidayattullah, and W. Suteddy, “Studi Komparasi Algoritma SVM dan Random Forest Pada Analisis Sentimen Komentar YouTube BTS,” 2023.
- [11] R. Firdaus, R. Al Hariri, and H. F. Amran, “Sentimen Analisis Masyarakat Tentang Penetapan Hari Raya Idul Adha Tahun 2023 Pada Video Youtube Menggunakan

Algoritma Random Forest dan Support Vector Machine,” JURNAL FASILKOM, vol. 14, no. 1, pp. 278–285, 2024.