KLASIFIKASI SENTIMEN MASYARAKAT TERHADAP KEBIJAKAN EKONOMI "PURBAYA EFFECT" MENGGUNAKAN ALGORITMA SVM DAN RANDOM FOREST

SENTIMENT CLASSIFICATION OF PUBLIC OPINION ON THE "PURBAYA EFFECT" ECONOMIC POLICY USING SVM AND RANDOM FOREST ALGORITHMS

Lailatus Syarifah[1*], Abd. Ghofur[2], Lukman Faqih Lidimilah[3]

*E-mail: lailasyarifahjm@gmail.com

[1,2,3] Program Studi Teknologi Informasi, Universitas Ibrahimy

Abstrak

Kebijakan ekonomi Purbaya Effect menjadi salah satu isu yang mendapat perhatian publik karena berkaitan dengan stabilitas ekonomi nasional, khususnya pengendalian inflasi dan pengelolaan defisit anggaran. Analisis sentimen terhadap isu ini masih menghadapi tantangan akibat penggunaan istilah ekonomi yang spesifik dan karakteristik bahasa media sosial yang tidak baku. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat terhadap kebijakan Purbaya Effect dengan membandingkan algoritma Support Vector Machine (SVM) dan Random Forest. Dataset yang digunakan terdiri atas 3.792 komentar valid dari tiga kanal YouTube yang diproses melalui tahapan preprocessing dan pelabelan sentimen berbasis leksikon yang diperkaya terminologi ekonomi. Hasil pelabelan menunjukkan bahwa sentimen negatif mendominasi sebesar 56,3%, sedangkan sentimen positif sebesar 43,7%. Ekstraksi fitur dilakukan menggunakan TF-IDF dengan 2.000 fitur, sedangkan optimasi parameter menggunakan GridSearchCV. Hasil evaluasi menunjukkan bahwa SVM memperoleh performa terbaik dengan akurasi sebesar 87,35%, lebih tinggi dibandingkan Random Forest sebesar 86,30%. Dominasi sentimen negatif mengindikasikan adanya kekhawatiran masyarakat terhadap stabilitas ekonomi, pengelolaan fiskal, dan dampak kebijakan terhadap kondisi ekonomi. Hasil penelitian menunjukkan bahwa SVM lebih efektif dalam klasifikasi sentimen pada data teks kebijakan ekonomi serta dapat digunakan untuk mendukung pemahaman terhadap persepsi publik.

Kata kunci: *klasifikasi sentimen, kebijakan ekonomi, support vector machine, random forest, lexicon-based labeling*

Abstract

*The *Purbaya Effect* economic policy has attracted public attention due to its association with national economic stability, particularly in controlling inflation and managing the budget deficit. Sentiment analysis in this domain remains challenging because of the use of specialized economic terminology and the informal nature of language on social media. This study aims to classify public sentiment toward the *Purbaya Effect* policy by comparing the performance of Support Vector Machine (SVM) and Random Forest algorithms. The dataset consists of 3,792 valid comments collected from three YouTube channels, which were processed through*

preprocessing stages and lexicon-based sentiment labeling enriched with economic terminology. The labeling results indicate that negative sentiment dominates at 56.3%, while positive sentiment accounts for 43.7%. Feature extraction was performed using TF-IDF with 2,000 selected features, and parameter optimization was conducted using GridSearchCV. The evaluation results show that SVM achieved the best performance with an accuracy of 87.35%, outperforming Random Forest, which achieved an accuracy of 86.30%. The dominance of negative sentiment suggests public concerns regarding economic stability, fiscal management, and the potential impact of the policy on economic conditions. The findings demonstrate that SVM is more effective for sentiment classification on economic policy-related text and can support a better understanding of public perception toward government policies

Keywords: *sentiment classification, economic policy, support vector machine, random forest, lexicon-based labeling*

1. PENDAHULUAN

Perkembangan teknologi digital telah mengubah cara masyarakat menyampaikan opini terhadap kebijakan pemerintah, khususnya melalui *platform* media sosial. *YouTube* menjadi salah satu media yang aktif digunakan sebagai sarana diskusi publik, di mana pengguna dapat menyampaikan dukungan, kritik, maupun kekhawatiran terhadap isu-isu nasional, termasuk kebijakan ekonomi. Komentar yang dihasilkan pada platform ini merepresentasikan respons spontan masyarakat dan berpotensi dimanfaatkan sebagai sumber data untuk analisis sentimen secara *real-time* [1]. Dalam konteks pengolahan teks, analisis sentimen merupakan bagian dari *Natural Language Processing (NLP)* yang digunakan untuk mengidentifikasi kecenderungan emosi atau opini dalam suatu teks [2].

Penerapan metode ini menjadi relevan dalam menganalisis fenomena “Purbaya Effect” sebagai kebijakan fiskal ekspansif yang memicu diskusi luas di media sosial dan menghasilkan berbagai opini dari masyarakat [3], [4]. Meskipun penelitian sebelumnya telah dilakukan menggunakan *metode Logistic Regression* pada *platform Instagram* [5], penelitian tersebut masih memiliki keterbatasan dalam mengatasi subjektivitas pelabelan, kompleksitas data yang mengandung *noise,* serta keterbatasan kamus sentimen pada domain ekonomi.

Tantangan dalam analisis sentimen pada domain kebijakan ekonomi terletak pada penggunaan terminologi khusus yang tidak umum dalam kamus sentimen standar. Istilah seperti “likuiditas”, “fiskal ekspansif”, dan “IHSG” memiliki konteks makna yang berbeda dibandingkan dengan kosakata umum. Selain itu, sebagian besar penelitian NLP bahasa Indonesia masih berfokus pada domain hiburan, politik umum, atau ulasan produk, sehingga masih terdapat kesenjangan penelitian pada domain kebijakan ekonomi. Oleh karena itu, penelitian ini menawarkan pendekatan *lexicon-based labeling* yang diperkaya dengan terminologi ekonomi untuk meningkatkan kualitas pelabelan data.

Untuk melakukan klasifikasi sentimen, penelitian ini membandingkan algoritma *Support Vector Machine (SVM)* dan *Random Forest*. SVM dipilih karena kemampuannya dalam menangani data teks berdimensi tinggi, sedangkan *Random Forest* digunakan sebagai pembanding karena kemampuannya dalam menangani data yang mengandung *noise*[6], [7]. Berbeda dengan penelitian sebelumnya yang umumnya menggunakan kamus sentimen umum dan berfokus pada satu metode klasifikasi, penelitian ini mengintegrasikan pendekatan lexicon-based labeling yang diperkaya terminologi ekonomi serta membandingkan performa SVM dan Random Forest pada data komentar YouTube. Hasil penelitian diharapkan dapat memberikan gambaran mengenai persepsi masyarakat terhadap kebijakan Purbaya Effect

sekaligus menjadi referensi bagi pengembangan penelitian analisis sentimen pada domain kebijakan ekonomi

2. METODOLOGI PENELITIAN

Metodologi penelitian ini dirancang untuk menguraikan tahapan klasifikasi sentimen terhadap kebijakan ekonomi "Purbaya Effect" secara sistematis dan dapat direplikasi. Penelitian ini mengadopsi kerangka kerja *Knowledge Discovery in Database (KDD)* yang terdiri dari tahapan data *selection, preprocessing, transformation*, data mining, dan *evaluation* [8]. Kerangka ini digunakan untuk memastikan proses pengolahan data berjalan terstruktur mulai dari data mentah hingga menghasilkan model klasifikasi yang valid.

2.1 Pendekatan Penelitian

Penelitian ini menerapkan pendekatan kuantitatif dengan desain eksperimen komparatif. Melalui teknik *Natural Language Processing* (NLP), opini publik ditransformasikan secara digital untuk diklasifikasikan ke dalam kategori sentimen positif dan negatif menggunakan algoritma *Support Vector Machine (SVM)* dan *Random Forest*. Jenis penelitian yang digunakan *adalah text mining berbasis supervised learning* untuk menganalisis opini publik terhadap kebijakan fiskal melalui data komentar *YouTube.*

2.2 Jenis Penelitian

Jenis penelitian ini adalah klasifikasi sentimen berbasis *text mining* dengan landasan operasional KDD [8]. Studi ini difokuskan pada pengukuran persepsi masyarakat terhadap fenomena kebijakan fiskal melalui data tekstual yang bersumber dari platform *YouTube.*

2.3 Sumber dan Teknik Pengumpulan Data

a. Pengumpulan Data

Data primer diperoleh melalui *YouTube Data API v3* dengan mengekstraksi komentar dari tiga kanal, yaitu Raymond Chin, Metro TV, dan CNBC Indonesia. Ketiga kanal tersebut dipilih berdasarkan tingkat interaksi (*engagement*) serta relevansi yang tinggi terhadap isu "Purbaya Effect" Rentang waktu pengambilan data mencakup periode September 2025 hingga Februari 2026 guna merepresentasikan dinamika opini publik secara luas. Secara terperinci, data diambil dari video kanal Metro TV (4 Oktober 2025–16 Februari 2026), CNBC Indonesia (3 November 2025–24 November 2025), dan Raymond Chin (23 September 2025–6 November 2025). Meskipun setiap video memiliki fokus pembahasan yang berbeda, seluruhnya tetap berada dalam konteks kebijakan ekonomi yang sama. Setelah dilakukan proses filtrasi untuk memastikan validitas dan relevansi data, diperoleh 3.792 data valid hasil preprocessing dari total 3.833 komentar yang akan digunakan dalam tahap klasifikasi.

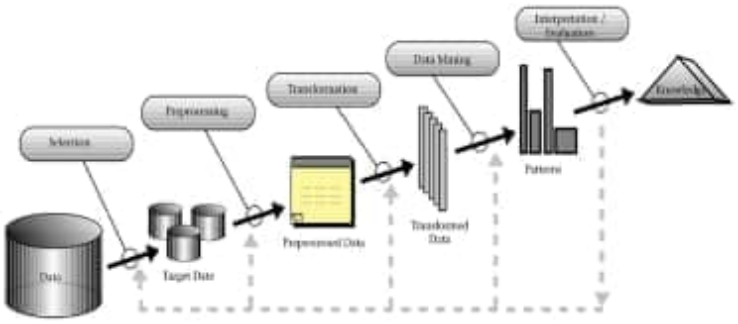
b. Kajian Pustaka

Kajian ini berfungsi memperkuat landasan teoretis mengenai arsitektur algoritma SVM, mekanisme *Random Forest*, serta standardisasi pengolahan teks bahasa Indonesia

2.4 Metode Pengembangan Sistem

Mengacu pada siklus KDD, tahapan analisis data dilakukan melalui prosedur berikut [8]:

a. ***Text Preprocessing***: Tahap ini meliputi cleaning, case folding, tokenizing, normalization, negation handling, stopword removal, dan stemming menggunakan Sastrawi untuk menghasilkan data teks yang lebih bersih dan terstruktur sebelum proses klasifikasi. Selain itu, dilakukan normalisasi terminologi ekonomi untuk menyeragamkan istilah khusus, seperti "ihsg" menjadi "indeks harga saham gabungan".
b. **Pelabelan Leksikon**: Penentuan label dilakukan secara otomatis menggunakan pendekatan lexicon-based labeling dengan kamus sentimen yang terdiri dari 139 kata positif dan 202 kata negatif yang telah disesuaikan dengan konteks ekonomi dan kebijakan publik.



Gambar 1. Tahapan Metode Knowledge Discovery in Database (KDD)

c. **Representasi Fitur**: Menggunakan *TF-IDF Vectorizer* dengan parameter *ngram_range=(1,2)* untuk menangkap pola kata tunggal maupun frasa.
d. **Konstruksi Model**: Model SVM dan Random Forest dilatih menggunakan GridSearchCV untuk memperoleh kombinasi parameter terbaik. Selain itu, parameter class_weight='balanced' diterapkan untuk mengurangi potensi bias klasifikasi antar kelas.
e. **Evaluasi Model:** Kinerja model dievaluasi menggunakan metrik Accuracy, Precision, Recall, dan F1-Score berdasarkan data pengujian. Selain itu, confusion matrix digunakan untuk menganalisis kemampuan model dalam mengklasifikasikan sentimen positif dan negatif.

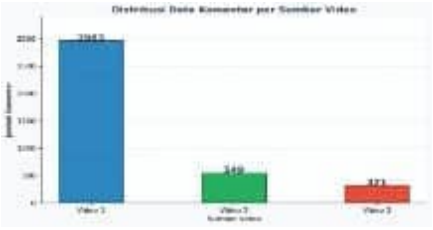
3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil penelitian klasifikasi sentimen terhadap kebijakan "Purbaya Effect". Proses penelitian dilakukan mulai dari ekstraksi data hingga interpretasi hasil untuk memastikan validitas temuan [9]. Alur penelitian ditunjukkan pada **Gambar 2.**

Gambar 2. Blok Diagram Tahapan Alur Proses Penelitian

3.1 Pengumpulan Data dan Penggabungan Data

Data penelitian diekstraksi dari *platform YouTube* menggunakan *YouTube Data API v3* melalui pustaka *google-api-python-client* [10]. Tiga kanal yang digunakan mewakili perspektif yang berbeda, yaitu Raymond Chin (edukasi ekonomi makro), Metro TV (analisis sosiopolitik), danCNBC Indonesia (analisis bisnis dan pasar). Total data yang berhasil dihimpun sebanyak3.833 komentar mentah. Distribusi data dari masing-masing sumber ditampilkan pada **Gambar 3.**



Gambar 3. Distribusi Data Komentar per Sumber Video

3.2 *Preprocessing Teks*

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data dengan menghilangkan *(noise)* serta melakukan standarisasi pada teks agar dapat diproses lebih lanjut oleh algoritma klasifikasi [11]. Setiap tahapan *preprocessing* dijelaskan sebagai berikut:

Tabel 1. Tahapan dan Hasil Preprocessing Teks

Tahap	Deskripsi	Contoh Input	Contoh Output
Cleaning	Hapus URL, emoji, mention,hashtag, angka	Kebijakan ini mantap!! 👍 100% bantu UMKM. @bi_gov #ekonomi	Kebijakan ini mantap bantu UMKM ekonomi
Case Folding	Ubah huruf menjadi lowercase	Kebijakan ini mantap bantu UMKM ekonomi	kebijakan ini mantap bantu umkm ekonomi
Tokenizing	Memecah kalimat menjadi token atau kata	kebijakan ini mantap bantu umkm ekonomi	['kebijakan', 'ini', 'mantap', 'bantu', 'umkm', 'ekonomi']
Normalisasi	Ubah slang/istilah ekonomi	umkm	usaha mikro kecil menengah
Negation Handling	Menggabungkan kata negasi dengan kata setelahnya agar makna tidak berubah	tidak setuju dengan kebijakan ini	tidak_setuju dengan kebijakan ini
Stopword Removal	Hapus kata tidak penting	kebijakan ini mantap bantu usaha mikro kecil menengah ekonomi	kebijakan mantap bantu usaha mikro kecil menengah ekonomi
***Stemming* (Sastrawi)**	Ubah kata ke bentuk dasar	kebijakan mantap bantu usaha mikro kecil menengah ekonomi	bijak mantap bantu usaha mikro kecil menengah ekonomi

Tahapan *preprocessing* dilakukan secara berurutan untuk menghasilkan representasi teks yang lebih bersih dan konsisten sebelum proses ekstraksi fitur. Setelah melalui proses validasi, diperoleh **3.792 data valid** dari total 3.833 data awal.

3.3 Pelabelan Berbasis Leksikon

Pelabelan sentimen pada penelitian ini dilakukan secara otomatis menggunakan pendekatan lexicon-based labeling dengan kamus sentimen yang terdiri dari 139 kata positif dan 202 kata negatif yang telah diperkaya dengan terminologi ekonomi. Proses pelabelan dilakukan dengan menghitung skor sentimen berdasarkan kemunculan kata-kata yang terdapat pada hasil preprocessing. Suatu komentar diklasifikasikan sebagai sentimen positif apabila jumlah skor positif lebih besar daripada skor negatif, sedangkan komentar diklasifikasikan sebagai sentimen negatif apabila jumlah skor negatif lebih dominan[12]. Hasil pelabelan menunjukkan bahwa dari 3.792 komentar valid, sebanyak 2.134 komentar (56,3%) termasuk ke dalam sentimen negatif dan 1.658 komentar (43,7%) termasuk ke dalam sentimen positif. Selain itu, terdapat 1.248 komentar yang memiliki skor sentimen netral (skor = 0) sehingga memerlukan mekanisme tiebreak untuk menentukan kelas akhir. Distribusi sentimen menunjukkan bahwa persepsi masyarakat terhadap kebijakan Purbaya Effect cenderung didominasi oleh sentimen negatif. Kondisi ini mengindikasikan adanya kekhawatiran masyarakat terhadap berbagai aspek kebijakan ekonomi, seperti stabilitas ekonomi, pengelolaan fiskal, dan dampak kebijakan terhadap kondisi ekonomi masyarakat. Distribusi hasil pelabelan sentimen ditunjukkan pada **Gambar 4.**

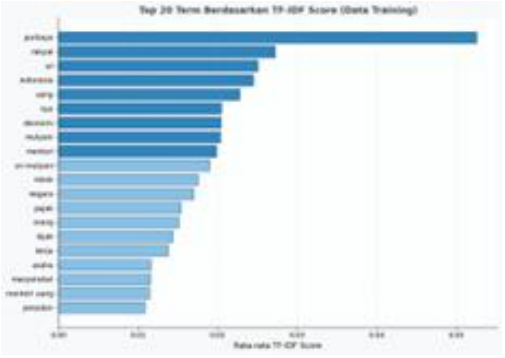
Gambar 4. Hasil *Labeling Lexicon-Based* dan Proporsi Kelas

3.4 Pembagian Data dan Ekstraksi Fitur

Data dibagi menggunakan *stratified train-test split* dengan rasio 80:20 untuk memastikan proporsi kelas tetap terjaga secara konsisten pada kedua subset data [13]. Proses ini menghasilkan 3.033 data latih *(training)* dan 759 data uji *(testing)* dengan *(random_state=42).* Distribusi proporsi kelas pada hasil pembagian data disajikan pada **Gambar 5.**

Gambar 5. Distribusi Kelas pada Data *Training* dan *Testing*

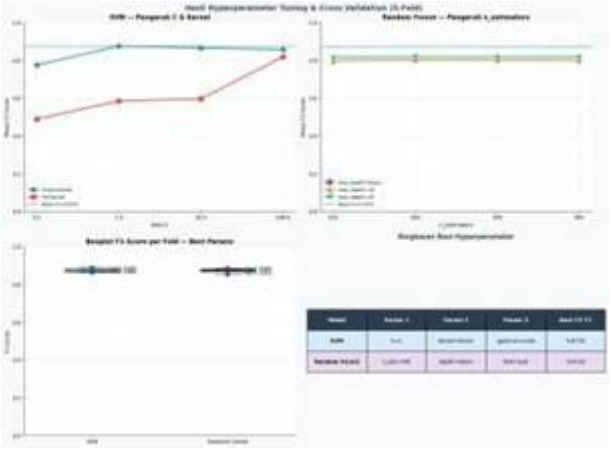
Teks diubah menjadi representasi numerik menggunakan *TF-IDF Vectorizer* [14], dengan konfigurasi: *max_features=2.000, ngram_range=(1,2), min_df=2,* dan *sublinear_tf=True*. Penggunaan bigram *(ngram_range=(1,2))* memungkinkan model menangkap konteks frasa penting seperti "ekonomi" atau "menteri uang" yang memiliki makna lebih spesifik dibandingkan kata tunggal *(unigram)*. Hasil ekstraksi fitur beserta 20 kata dengan bobot TF-IDF tertinggi ditampilkan pada **Gambar 6**.



Gambar 6. Top 20 *Term* Berdasarkan *TF-IDF Score*

3.5 Hasil *Hyperparameter Tuning*

Proses optimasi parameter menggunakan *GridSearchCV* menghasilkan konfigurasi terbaik yang berbeda untuk masing-masing model [15].



Gambar 7. Hasil *Hyperparameter Tuning* dan *Cross Validation (5-Fold)*.

Temuan ini mengonfirmasi bahwa data teks setelah transformasi TF-IDF bersifat *linearly separable* di ruang fitur berdimensi tinggi (2.000 dimensi), sehingga pemisahan linear lebih efektif dalam menentukan *hyperplane* optimal [16]. Ringkasan parameter terbaik SVM dengan C = 1, kernel = linear, gamma = scale, dan class_weight = balanced dengan nilai F1-Score sebesar 0,8739. Sementara itu, Random Forest memperoleh konfigurasi terbaik pada parameter n_estimators = 300, max_depth = None, min_samples_split = 2, max_features = sqrt, dan class_weight = balanced dengan nilai F1-Score sebesar 0,8720. Hasil ini menunjukkan bahwa SVM sedikit lebih unggul dibandingkan Random Forest dalam mengklasifikasikan data sentimen berbasis TF-IDF, yang mengindikasikan bahwa data dapat dipisahkan secara efektif menggunakan kernel linear.

3.6 Evaluasi Model

Evaluasi model dilakukan pada 759 data uji menggunakan metrik Accuracy, Precision, Recall, dan F1-Score untuk mengukur performa klasifikasi sentimen. Penggunaan beberapa metrik ini bertujuan memberikan gambaran yang lebih komprehensif mengenai kemampuan model dalam mengklasifikasikan sentimen positif dan [17]. Nilai Precision, Recall, dan F1-Score dihitung menggunakan pendekatan macro average, sehingga setiap kelas memperoleh bobot yang sama dalam proses evaluasi. Pendekatan ini dipilih untuk memberikan penilaian yang lebih seimbang terhadap performa model pada masing-masing kelas sentimen. Hasil evaluasi kedua model ditunjukkan pada Tabel 2.

Tabel 2. Evaluasi Model

Algoritma	***Accuracy***	***Precision***	***Recall***	***F-1 Score***
SVM	0,8735%	0,8719%	0,8708%	0,8713%
Random Forest	0,8630%	0,8604%	0,8645%	0,8617%

Berdasarkan Tabel 2, model SVM memperoleh performa terbaik dengan nilai Accuracy sebesar 0,8735 dan F1-Score sebesar 0,8713. Sementara itu, Random Forest memperoleh Accuracy sebesar 0,8630 dan F1-Score sebesar 0,8617. Hasil tersebut menunjukkan bahwa kedua model mampu melakukan klasifikasi sentimen dengan baik, namun SVM memberikan performa yang lebih unggul secara keseluruhan. Perbandingan confusion matrix kedua model ditunjukkan pada Gambar 8.

Gambar 8. Perbandingan *Confusion Matrix Svm dan Random Forest*

Berdasarkan gambar 8 perbandingan confusion matrix menunjukkan bahwa model SVM mampu mengklasifikasikan 663 komentar dengan benar dari total 759 data uji, terdiri atas 282 True Positive (TP) dan 381 True Negative (TN). Sementara itu, Random Forest berhasil mengklasifikasikan 655 komentar dengan benar, terdiri atas 291 True Positive (TP) dan 364 True Negative (TN). Hasil ini menunjukkan bahwa SVM memiliki kemampuan yang lebih baik

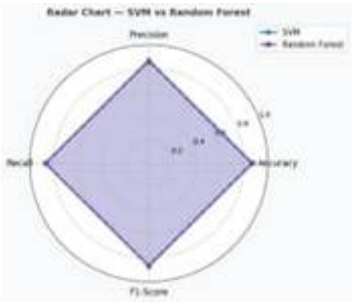
dalam membedakan sentimen positif dan negatif pada data komentar terkait kebijakan Purbaya Effect.

3.7 Analisis Perbandingan Model

Analisis perbandingan model dilakukan untuk mengevaluasi kinerja Support Vector Machine (SVM) dan Random Forest berdasarkan metrik Accuracy, Precision, Recall, dan F1-Score. Berdasarkan hasil evaluasi pada bar chart **Gambar 9**, kedua model menunjukkan performa yang baik dalam melakukan klasifikasi sentimen. Namun, SVM memperoleh nilai Accuracy sebesar 0,8735 dan F1-Score sebesar 0,8713, sedikit lebih tinggi dibandingkan Random Forest yang memperoleh Accuracy sebesar 0,8630 dan F1-Score sebesar 0,8617.

Gambar 9. Hasil Perbandingan Model

Perbandingan performa kedua model juga divisualisasikan melalui radar chart pada **Gambar 9.** Visualisasi tersebut menunjukkan bahwa SVM memiliki performa yang lebih stabil dan unggul pada sebagian besar metrik evaluasi. Selain itu, hasil cross-validation pada tahap hyperparameter tuning juga menunjukkan bahwa SVM memperoleh nilai F1-Score validasi yang lebih tinggi dibandingkan Random Forest, yaitu 0,8739 dan 0,8720.



Gambar 10. Radar Chart

Keunggulan SVM pada penelitian ini diduga berkaitan dengan kemampuannya dalam menangani data teks berdimensi tinggi hasil transformasi TF-IDF. Penggunaan kernel linear dengan parameter C = 1 memungkinkan model membentuk hyperplane yang optimal untuk memisahkan kelas sentimen secara efektif [18], Oleh karena itu, SVM dapat dianggap sebagai model yang lebih sesuai untuk diimplementasikan pada sistem klasifikasi sentimen terkait kebijakan ekonomi Purbaya Effect.Keunggulan SVM berkaitan dengan kemampuan kernel linear dalam memisahkan data berdimensi tinggi secara efektif .

4. KESIMPULAN DAN SARAN

Hasil penelitian menunjukkan bahwa algoritma SVM (Support Vector Machine) memberikan performa lebih baik dibandingkan Random Forest dalam klasifikasi sentimen

terhadap kebijakan Purbaya Effect dengan nilai Accuracy 0,8735 dan F1-Score 0,8713 berbanding 0,8630 dan 0,8617. Hal ini membuktikan TF-IDF efektif dalam representasi fitur numerik untuk data teks. Dari proses pelabelan menggunakan metode lexicon-based dengan terminologi ekonomi, diperoleh distribusi sentimen 56,3% negatif dan 43,7% positif, yang mana dominasi sentimen negatif tersebut mencerminkan kekhawatiran masyarakat terhadap berbagai aspek kebijakan ekonomi, seperti kenaikan inflasi yang meningkat, ketidakpastian stabilitas ekonomi nasional, pengelolaan fiskal yang kurang transparan, serta dampak kebijakan terhadap daya beli dan kesejahteraan masyarakat. Untuk pengembangan penelitian selanjutnya, disarankan menggunakan pendekatan deep learning seperti LSTM, Bidirectional LSTM, atau IndoBERT untuk membandingkan performa dengan metode machine learning konvensional, melakukan validasi pelabelan manual atau oleh beberapa annotator guna meningkatkan kualitas label, serta memanfaatkan sumber data dari berbagai platform seperti X (Twitter), Instagram, atau portal berita daring agar dapat memberikan gambaran persepsi publik yang lebih komprehensif terhadap kebijakan ekonomi.

5. DAFTAR RUJUKAN

[1] L. Septiana, V. Yasin, and A. Z. Sianipar, "*Analyzing public sentiment on YouTube comments regarding the free lunch policy using the Support Vector Machine (SVM) algorithm*," *Jurnal Mandiri IT*, vol. 14, no. 1, pp. 130–137, 2025.

[2] I. Agusta and A. Rahman, "Penerapan *natural language processing* untuk analisis sentimen terhadap kebijakan pemerintah," *Jurnal Kebangsaan RI*, vol. 1, no. 1, pp. 45–56, 2025.

[3] Kementerian Keuangan Republik Indonesia, *Laporan Kebijakan Fiskal 2025: Penguatan Peran Fiskal dalam Akselerasi Pertumbuhan Ekonomi*. Jakarta: Direktorat Jenderal Anggaran, 2025.

[4] Kementerian Keuangan Republik Indonesia, *Peraturan Menteri Keuangan Nomor 63 Tahun 2025 tentang Pemanfaatan Saldo Anggaran Lebih untuk Penguatan Likuiditas Perbankan Nasional*. Jakarta, 2025.

[5] A. Khaidar, "Analisis sentimen di *Instagram* terhadap Menteri Keuangan Purbaya Yudhi Sadewa menggunakan metode logistic regression," *JITET (Jurnal Informatika dan Teknik Elektro Terapan)*, vol. 13, no. 2, pp. 1674–1683, 2025.

[6] A. Nurfaza and S. T. Salim, "Klasifikasi *cyberbullying* pada komentar video *YouTube* menggunakan metode random forest," *Jurnal Ilmiah Teknologi dan Informatika*, vol. 9, no. 1, pp. 45–54, 2023.

[7] Hidayat, F. Santoso, and L. F. Lidimillah, "Analisis sentimen pengguna *YouTube* tentang Rohingya menggunakan algoritma *Support Vector Machine*," *G-Tech: Jurnal Teknologi Terapan*, vol. 8, no. 4, pp. 1729–1738, 2024.

[8] R. Swastika et al., Implementasi Data Mining (*Clustering, Association, Prediction, Estimation, Classification*). Lampung: CV. Adanu Abimata, 2023.

[9] H. S. Mulyono and U. Saprudin, "Efektivitas *logistic regression* dalam analisis sentimen berbahasa Indonesia pada komentar *YouTube* tentang isu ketenagakerjaan," *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, vol. 6, no. 3, pp. 1547–1555, 2025.

[10] N. Huwaida, "Analisis sentimen komentar *YouTube* terhadap pemindahan ibu kota negara menggunakan metode Naïve Bayes," *Jambura Journal of Informatics*, vol. 6, no. 1, 2024.

[11] N. Arifin, U. Enri, and N. Sulistiyowati, "Penerapan algoritma *Support Vector Machine (SVM)* dengan *TF-IDF n-gram* untuk *text classification*," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 6, no. 2, pp. 129–136, 2021.

[12] W. F. Abdillah, A. Premana, and R. M. H. Bhakti, "Analisis sentimen penanganan COVID-19 dengan *Support Vector Machine*: evaluasi leksikon dan metode ekstraksi fitur," *Jurnal Ilmiah Intech: Information Technology Journal of UMUS*, vol. 3, no. 2, pp. 160–170, 2021.

[13] Y. I. Muasaroh et al., "Analisis sentimen komentar *YouTube* terhadap isu ijazah Presiden Jokowi menggunakan *Support Vector Machine* dan *Random Forest*," in *Prosiding Semnas Sekolah Tinggi Teknologi Dumai*, 2025, pp. 371–380.

[14] S. W. Iriananda, R. W. Budiawan, and A. Y. Rahman, "Optimasi klasifikasi sentimen komentar pengguna game bergerak menggunakan SVM, *grid search*, dan *kombinasi n-gram*," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, pp. 743–752, 2024.

[15] T. Misriati and R. Aryanti, "Optimalisasi *Random Forest* dan *Support Vector Machine* dengan *hyperparameter GridSearchCV* untuk analisis sentimen ulasan PrimaKu," *Journal of Information System Research (JOSH)*, vol. 5, no. 4, pp. 1333–1341, 2024.

[16] M. Fajri and A. Primajaya, "Komparasi teknik *hyperparameter optimization* pada SVM menggunakan *grid search* dan *random search*," *Journal of Applied Informatics and Computing (JAIC)*, vol. 7, no. 1, pp. 14–19, 2023.

[17] F. N. Hasan *et al.*, "Analisis perbandingan algoritma SVM dan *Random Forest* untuk klasifikasi sentimen pada data teks berdimensi tinggi," *Jurnal Media Informatika Budidarma*, vol. 7, no. 3, pp. 517–523, 2023.

[18] A. P. Agusta *et al.*, "Analisis sentimen kebijakan publik menggunakan *natural language processing:* studi kasus opini masyarakat terhadap isu ekonomi," *Jurnal Informatika dan Sistem Informasi*, vol. 5, no. 1, pp. 88–97, 2024.