

Pengembangan Model Klasifikasi Belanja Pegawai Daerah Tahun Anggaran 2025-2026 Menggunakan Algoritma *Support Vector Machine*

Tsabit Faishal Rasyid¹, Imam Budiawan^{2*}, Yumi Novita Dewi³

^{1,2,3}Informatika, Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, Kec. Senen, Kota Jakarta Pusat, DKI Jakarta
imam.imb@bsi.ac.id

Abstract

Inefficient employee expenditure allocation remains a critical issue in Indonesian regional governments, with most regions allocating over 30% of their budget to personnel costs. This study develops a classification model using Support Vector Machine (SVM) to categorize the efficiency level of employee expenditure across 546 regional governments using data from the State Financial Information System (SIPD) for fiscal years 2025-2026. Data preprocessing included Min-Max normalization, feature engineering, and stratified 80:20 train-test splitting. Hyperparameter optimization through Grid Search with 5-Fold Stratified Cross Validation identified the optimal configuration: RBF kernel, C=100, gamma=1. The model achieved accuracy of 98.63%, precision of 98.62%, recall of 98.63%, and F1-Score of 98.62% on the test set. Feature importance analysis revealed that the employee expenditure ratio is the most dominant predictor. The model correctly identified 87.4% of regions in 2026 as inefficient, providing actionable insights for fiscal policy reform.

Keywords: Support Vector Machine, employee expenditure, classification, regional finance, efficiency

Abstrak

Ketidakefisienan alokasi belanja pegawai menjadi permasalahan kritis di hampir seluruh pemerintah daerah Indonesia, dengan sebagian besar daerah mengalokasikan lebih dari 30% anggaran untuk komponen personnel. Penelitian ini mengembangkan model klasifikasi menggunakan Support Vector Machine (SVM) untuk mengelompokkan tingkat efisiensi belanja pegawai pada 546 pemerintah daerah berdasarkan data Sistem Informasi Pemerintahan Daerah (SIPD) periode Tahun Anggaran 2025-2026. Pra-pemrosesan data mencakup normalisasi Min-Max, rekayasa fitur, dan pembagian data terstratifikasi 80:20. Optimasi hyperparameter melalui Grid Search dengan 5-Fold Stratified Cross Validation menghasilkan konfigurasi optimal: kernel RBF, C=100, gamma=1. Model mencapai akurasi 98,63%, presisi 98,62%, recall 98,63%, dan F1-Score 98,62% pada data pengujian. Analisis kepentingan fitur menunjukkan bahwa rasio belanja pegawai merupakan prediktor paling dominan. Model secara akurat mengidentifikasi 87,4% daerah pada TA 2026 termasuk kategori tidak efisien, memberikan landasan data yang kuat bagi reformasi kebijakan fiskal daerah.

Kata Kunci: Support Vector Machine, belanja pegawai, klasifikasi, keuangan daerah, efisiensi

Copyright (c) 2026 Tsabit Faishal Rasyid, Imam Budiawan, Yumi Novita Dewi

✉ Corresponding author: Tsabit Faishal Rasyid

Email Address: imam.imb@bsi.ac.id (Jl. Kramat Raya No.98, Kota Jakarta Pusat, DKI Jakarta)

Received 01 August 2026, Accepted 13 August 2026, Published 26 August 2026

PENDAHULUAN

Pengelolaan keuangan daerah merupakan aspek kritis dalam tata kelola pemerintahan Indonesia. Salah satu komponen terbesar dalam Anggaran Pendapatan dan Belanja Daerah (APBD) adalah belanja pegawai, yang mencakup gaji pokok, tunjangan, dan komponen remunerasi lainnya. Ketidakefisienan alokasi belanja pegawai dapat mengurangi kapasitas fiskal untuk membiayai sektor-sektor produktif seperti infrastruktur, pendidikan, dan layanan publik. Data menunjukkan bahwa proporsi belanja pegawai di berbagai daerah sangat bervariasi, dari di bawah 40% hingga mencapai 60-70% dari total APBD [1].

Pemerintah pusat telah menerbitkan Instruksi Presiden Nomor 1 Tahun 2025 tentang efisiensi anggaran, yang mempertegas tekanan untuk mengurangi belanja yang dinilai tidak strategis, termasuk belanja pegawai yang berlebihan. Dalam konteks ini, kemampuan untuk mengklasifikasikan daerah mana

yang sudah efisien dan mana yang belum menjadi sangat penting. Namun, proses evaluasi terhadap efisiensi belanja pegawai masih banyak dilakukan secara manual, berbasis intuisi, atau dengan perbandingan rasio sederhana [2].

Machine learning menawarkan solusi yang lebih objektif dan terukur. Berbagai penelitian telah membuktikan bahwa Support Vector Machine (SVM) mampu mengklasifikasikan data dengan akurasi tinggi, bahkan pada dataset dengan banyak variabel. Yazid et al. [3] menggunakan SVM untuk mengklasifikasikan kelayakan penerima bantuan sosial APBD dengan akurasi testing 95%. Moningkey et al. [4] berhasil mengimplementasikan SVM untuk klasifikasi multiclass pada Indeks Kepuasan Masyarakat. Puspandoyo et al. [5] menunjukkan bahwa SVM menghasilkan akurasi tertinggi sebesar 96,1% dalam memprediksi kualitas laporan keuangan kementerian.

Namun, penelitian-penelitian terdahulu yang menerapkan machine learning di domain keuangan publik Indonesia mayoritas berfokus pada prediksi penerimaan pajak, deteksi fraud anggaran, atau evaluasi kinerja secara umum. Klasifikasi khusus terhadap efisiensi belanja pegawai daerah sebagai objek utama penelitian, apalagi dengan menggunakan pendekatan SVM yang terstruktur, belum banyak dieksplorasi. Celah inilah yang menjadi research gap dari penelitian ini.

Berdasarkan latar belakang tersebut, penelitian ini bermaksud membangun model klasifikasi berbasis SVM yang mampu mengelompokkan tingkat efisiensi belanja pegawai daerah berdasarkan data keuangan periode 2025-2026. Secara spesifik, tujuan penelitian ini meliputi: (1) mengidentifikasi faktor-faktor yang signifikan terhadap tingkat efisiensi belanja pegawai daerah, (2) membangun model klasifikasi berbasis SVM, (3) mengevaluasi performa model menggunakan metrik akurasi, presisi, recall, dan F1-Score, serta (4) menghasilkan rekomendasi berbasis hasil klasifikasi sebagai bahan masukan kebijakan pengelolaan belanja pegawai daerah.

Urgensi penelitian ini turut diperkuat oleh target penghematan anggaran nasional sebesar Rp306,7 triliun sebagaimana diamanatkan Instruksi Presiden Nomor 1 Tahun 2025, yang turut menysar pos belanja pegawai sebagai salah satu komponen yang perlu direstrukturasasi [1]. Tanpa instrumen klasifikasi yang objektif dan berbasis data, proses identifikasi daerah yang belanja pegawainya masih perlu dievaluasi akan terus berjalan lambat dan rentan terhadap bias subjektif.

METODE

Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif berbasis komputasi dengan menerapkan algoritma Support Vector Machine untuk mengklasifikasikan tingkat efisiensi belanja pegawai daerah. Data bersifat cross-sectional dua periode (2025 dan 2026), sehingga analisis dapat dilakukan baik secara per tahun maupun secara gabungan untuk menangkap pola yang lebih luas. Alur penelitian terdiri atas tahapan: pengumpulan data, eksplorasi data, preprocessing, rekayasa fitur, pembangunan model SVM, evaluasi model, dan interpretasi hasil [11].

Tahapan penelitian yang dilaksanakan secara berurutan dirangkum pada Tabel 1.

Tabel 1. Tahapan dan Kerangka Kerja Penelitian

No.	Tahapan	Keterangan
1	Pengumpulan Data	Mengunduh data rekap belanja pegawai daerah TA 2025-2026 dari SIPD Ditjen Bina Keuangan Daerah, mencakup 546 daerah di Indonesia.
2	Eksplorasi dan Pemahaman Data	Menganalisis struktur data, distribusi nilai, serta mengidentifikasi variabel yang relevan sebagai fitur model klasifikasi.
3	Preprocessing Data	Melakukan pembersihan data, normalisasi nilai numerik, penanganan missing value, serta pelabelan kelas efisiensi berdasarkan ambang batas rasio.
4	Rekayasa Fitur	Memilih dan membangun fitur paling informatif untuk model SVM, termasuk rasio belanja pegawai terhadap total belanja daerah.
5	Pembangunan Model SVM	Melatih model SVM menggunakan data training dengan berbagai konfigurasi kernel dan parameter.
6	Evaluasi Model	Mengukur performa model menggunakan metrik akurasi, presisi, recall, dan F1-Score pada data testing.
7	Interpretasi dan Rekomendasi	Menginterpretasikan hasil klasifikasi dan menyusun rekomendasi kebijakan berdasarkan pola efisiensi yang ditemukan.

Sumber: penelitian ini

Sumber Data

Data yang digunakan bersumber dari Sistem Informasi Pemerintahan Daerah (SIPD) Direktorat Jenderal Bina Keuangan Daerah, Kementerian Dalam Negeri Republik Indonesia. Data mencakup rekapitulasi anggaran belanja seluruh pemerintah daerah di Indonesia (tingkat Provinsi, Kabupaten, dan Kota) untuk dua tahun anggaran, yaitu TA 2025 dan TA 2026, dengan total 546 daerah pada masing-masing tahun. Total data gabungan (pooling data) yang digunakan adalah sebanyak 1.092 baris observasi. Tidak ditemukan missing values pada fitur-fitur keuangan utama, sehingga seluruh sampel dapat digunakan sepenuhnya dalam proses pemodelan[12].

Variabel-variabel keuangan utama yang tersedia dalam dataset ditunjukkan pada Tabel 2.

Tabel 2. Variabel Data Penelitian Utama

No.	Variabel	Keterangan
1	Belanja Daerah (Rp)	Total belanja daerah yang dianggarkan dalam APBD masing-masing daerah.
2	Belanja Pegawai Total (Rp)	Total belanja pegawai keseluruhan, termasuk tunjangan profesi dan tambahan penghasilan guru PNSD.
3	Belanja Pegawai Murni (Rp)	Belanja pegawai di luar Tunjangan Profesi Guru (TPG), Tunjangan Khusus Guru (TKG), dan Tambahan Penghasilan (Tamsil) Guru PNSD.
4	% Belanja Pegawai Murni terhadap Belanja Daerah	Rasio persentase belanja pegawai murni terhadap total belanja daerah; menjadi acuan utama pembentukan label kelas efisiensi.
5	Tahun Anggaran	Variabel penanda periode data (2025 atau 2026).

Sumber: penelitian ini

Preprocessing dan Rekayasa Fitur

Pembentukan label kelas efisiensi dilakukan berdasarkan rasio belanja pegawai murni (di luar TPG, TKG, Tamsil) terhadap total belanja daerah. Mengacu pada Undang-Undang Nomor 1 Tahun 2022 tentang Hubungan Keuangan antara Pemerintah Pusat dan Pemerintahan Daerah, batas maksimal porsi

belanja pegawai ditetapkan sebesar 30%. Klasifikasi yang digunakan adalah: Efisien (rasio $< 30\%$) dan Tidak Efisien (rasio $\geq 30\%$). Pembagian dataset dilakukan secara stratified random split dengan proporsi 80:20. Normalisasi dilakukan menggunakan MinMaxScaler ke rentang $[0, 1]$.

Implementasi SVM

Implementasi SVM dilakukan menggunakan library scikit-learn pada bahasa pemrograman Python. Optimasi hyperparameter dilakukan melalui Grid Search Cross Validation dengan 5-Fold Stratified Cross Validation. Parameter yang diuji meliputi kernel (Linear, RBF, Polynomial), regularization C (0,1; 1; 10; 100), gamma (0,001; 0,01; 0,1; 1), dan degree (2; 3; 4) untuk kernel polynomial [9]. Evaluasi model menggunakan empat metrik: akurasi, presisi, recall, dan F1-Score[13].

Alat dan Lingkungan Pengembangan

Seluruh proses analisis dan pemodelan pada penelitian ini dilaksanakan menggunakan perangkat lunak dan pustaka sebagaimana dirangkum pada Tabel 3.

Tabel 3. Alat dan Lingkungan Pengembangan Sistem

Komponen	Teknologi/Alat	Keterangan
Bahasa Pemrograman	Python	Versi 3.10 atau lebih tinggi
Preprocessing & Analisis	pandas, NumPy	Manipulasi dataframe dan komputasi numerik
Implementasi SVM	scikit-learn	Modul SVC dan GridSearchCV untuk optimasi parameter
Visualisasi	matplotlib, seaborn	Visualisasi distribusi data dan confusion matrix
Lingkungan Kerja	Jupyter Notebook / Google Colab	Notebook interaktif untuk eksplorasi dan dokumentasi kode

HASIL DAN DISKUSI

Statistik Deskriptif Data

Data yang digunakan terdiri dari 1.092 baris observasi (546 daerah x 2 tahun anggaran). Statistik deskriptif menunjukkan kesenjangan kemampuan keuangan yang sangat besar antar-daerah. Total belanja daerah berkisar dari Rp417,29 miliar hingga Rp82,66 triliun. Rasio Belanja Pegawai Murni memiliki nilai minimum 9,55% dan maksimum 61,39%, dengan median 37,54%, yang menunjukkan bahwa sebagian besar pemerintah daerah mengalokasikan lebih dari sepertiga anggarannya untuk belanja pegawai [5].

Rincian sebaran nilai untuk setiap variabel keuangan utama disajikan pada Tabel 4.

Tabel 4. Statistik Deskriptif Data Keuangan Daerah (n = 1.092)

Metrik Statistik	Belanja Daerah (Rp)	BP Murni (Rp)	BP Total (Rp)	Rasio BP Murni (%)	Rasio BP Total (%)
Minimum	417.294.709.442	150.655.943.391	169.758.576.391	9,55%	10,75%
Persentil 25%	1.032.765.204.652	393.847.408.420	448.905.562.554	32,71%	36,92%
Median (50%)	1.481.762.151.403	559.105.700.072	654.470.920.753	37,54%	42,94%
Persentil 75%	2.541.127.585.182	889.385.978.860	1.037.609.660.025	42,23%	47,81%

Maximum	82.663.062.617.784	21.276.227.503.102	22.316.492.000.507	61,39%	63,12%
---------	--------------------	--------------------	--------------------	--------	--------

Sumber: penelitian ini

Distribusi Kelas Aktual

Berdasarkan ambang batas 30% sesuai UU HKPD, distribusi kelas aktual menunjukkan dominasi kelas Tidak Efisien. Dari total 1.092 data gabungan, sebanyak 916 daerah (83,88%) melampaui ambang batas wajar. Pada TA 2025, 109 daerah (19,96%) terklasifikasi Efisien, sementara pada TA 2026 hanya 67 daerah (12,27%). Ketidakseimbangan kelas ini melatarbelakangi penggunaan teknik stratified split serta metrik F1-Score tertimbang pada tahap pemodelan [10].

Tabel 5. Distribusi Kelas Aktual Tingkat Efisiensi Belanja Pegawai

Tahun Anggaran	Efisien (< 30%)	Tidak Efisien (>= 30%)	Total	% Efisien
2025	109	437	546	19,96%
2026	67	479	546	12,27%
Total Gabungan	176	916	1.092	16,12%

Sumber: penelitian ini

Hasil Optimasi Hyperparameter

Grid Search 5-Fold Cross Validation menghasilkan parameter optimal: kernel RBF, C=100, gamma=1, dengan skor F1-Weighted terbaik sebesar 98,86%. Rata-rata akurasi pada validasi silang mencapai 98,85% dengan standar deviasi yang sangat kecil (0,0036), yang mengindikasikan tidak adanya gejala overfitting maupun underfitting. Kombinasi parameter ini konsisten dengan temuan Prasetyo dan Maulana [7] yang menunjukkan bahwa kernel RBF dengan optimasi parameter mampu menghasilkan performa tinggi pada data keuangan multidimensi [9].

Sebelum proses pelatihan, seluruh data dinormalisasi menggunakan metode Min-Max Normalization ke skala [0, 1], kemudian dibagi dengan proporsi 80:20 secara stratified random split untuk menjaga representasi proporsi kelas pada masing-masing subset. Hasil pembagian menghasilkan data pelatihan (training set) sebanyak 873 sampel, terdiri atas 141 sampel Efisien dan 732 sampel Tidak Efisien, serta data pengujian (testing set) sebanyak 219 sampel, terdiri atas 35 sampel Efisien dan 184 sampel Tidak Efisien.

Stabilitas model pada tahap validasi silang turut dievaluasi melalui 5-Fold Stratified Cross Validation pada data pelatihan, dengan hasil dirangkum pada Tabel 6.

Tabel 6. Hasil Uji 5-Fold Cross Validation pada Data Pelatihan (n = 873)

Metrik Evaluasi	Rata-rata (Mean)	Standar Deviasi (Std)	Nilai Minimum	Nilai Maksimum
Akurasi	0,9885	0,0036	0,9828	0,9943
F1-Score (Weighted)	0,9886	0,0037	0,9826	0,9942
Presisi (Weighted)	0,9888	0,0037	0,9826	0,9943
Recall (Weighted)	0,9885	0,0036	0,9828	0,9943

Sumber: penelitian ini

Rata-rata akurasi sebesar 98,85% dengan standar deviasi yang sangat kecil (0,0036) membuktikan bahwa model SVM yang dibangun memiliki tingkat stabilitas yang sangat tinggi serta tidak mengalami kendala overfitting maupun underfitting pada lipatan data (folds) yang berbeda [2][9].

Evaluasi Performa Model

Evaluasi performa model pada data pengujian (219 sampel) menghasilkan metrik yang sangat tinggi. Akurasi keseluruhan mencapai 98,63%, presisi weighted 98,62%, recall weighted 98,63%, dan F1-Score weighted 98,62%. Pada tingkat kelas, F1-Score untuk kelas Efisien mencapai 95,65% dan untuk kelas Tidak Efisien mencapai 99,19%. Model hanya menghasilkan 3 kesalahan dari 219 prediksi (misclassification rate = 1,37%). Hasil ini sejalan dengan temuan Puspandoyo et al. [5] yang mencapai akurasi 96,1% dan Yazid et al. [3] yang memperoleh akurasi 95% pada domain data keuangan pemerintah [5][3].

Tabel 7. Metrik Evaluasi Model SVM pada Data Pengujian

Metrik	Weighted	Macro
Akurasi	98,63%	-
Presisi	98,62%	97,99%
Recall	98,63%	96,87%
F1-Score	98,62%	97,42%

Sumber: penelitian ini

Untuk melihat lebih rinci kemampuan model dalam membedakan tiap kelas, Tabel 8 menyajikan laporan klasifikasi (classification report) per kelas pada data pengujian.

Tabel 8. Laporan Klasifikasi (Classification Report) per Kelas

Kelas Klasifikasi	Presisi	Recall	F1-Score	Jumlah Sampel
Efisien	97,06%	94,29%	95,65%	35
Tidak Efisien	98,92%	99,46%	99,19%	184
Rata-rata Makro	97,99%	96,87%	97,42%	219
Rata-rata Tertimbang	98,62%	98,63%	98,62%	219

Sumber: penelitian ini

Model mampu membedakan kelas Tidak Efisien dengan performa mendekati sempurna (F1-Score 99,19% dari 184 sampel). Untuk kelas Efisien yang jumlah sampelnya jauh lebih sedikit (35 sampel), model tetap menghasilkan F1-Score yang tinggi yaitu 95,65% (presisi 97,06% dan recall 94,29%), menunjukkan bahwa strategi stratified split dan metrik tertimbang berhasil mengatasi ketidakseimbangan kelas pada data pengujian [10].

Sebaran kesalahan klasifikasi model secara lebih detail digambarkan melalui confusion matrix pada Tabel 9.

Tabel 9. Confusion Matrix Pengujian Model

Aktual / Prediksi	Prediksi: Efisien	Prediksi: Tidak Efisien	Total Aktual
Aktual: Efisien	33 (True Negative)	2 (False Positive)	35
Aktual: Tidak Efisien	1 (False Negative)	183 (True Positive)	184
Total Prediksi	34	185	219

Sumber: penelitian ini

Dari 35 daerah yang aktualnya Efisien, model berhasil memprediksi 33 daerah secara benar, dengan hanya 2 daerah yang salah diklasifikasikan sebagai Tidak Efisien. Dari 184 daerah yang aktualnya Tidak Efisien, model berhasil memprediksi 183 daerah secara benar, dengan hanya 1 daerah yang keliru diprediksi sebagai Efisien. Tingkat kesalahan klasifikasi yang sangat rendah ini

(misclassification rate 1,37%) membuktikan keandalan algoritma SVM dengan kernel RBF dalam memetakan batas margin keputusan antar-kelas efisiensi belanja pegawai daerah [7][9].

Analisis Kepentingan Fitur

Analisis Permutation Feature Importance mengidentifikasi bahwa Rasio Belanja Pegawai Murni (%) merupakan fitur yang paling krusial terhadap performa prediksi model, dengan rata-rata penurunan skor F1 mencapai 0,1977. Secara berurutan, faktor-faktor lain yang turut berkontribusi adalah Rasio Belanja Pegawai Total (0,0675), Selisih Belanja Pegawai (0,0149), nominal Belanja Pegawai Total (0,0032), nominal Belanja Pegawai Murni (0,0028), variabel Tahun Anggaran (0,0021), dan nominal Belanja Daerah keseluruhan (0,0019). Temuan ini menunjukkan bahwa model tidak sekadar memisahkan kelas berdasarkan satu ambang batas tunggal, melainkan memanfaatkan ruang multidimensi non-linear melalui konstruksi hyperplane yang optimal[14].

Tabel 10. Nilai Kepentingan Fitur (Permutation Feature Importance)

Peringkat	Indikator Keuangan	Rata-rata Penurunan F1
1	Rasio Belanja Pegawai Murni (%)	0,1977
2	Rasio Belanja Pegawai Total (%)	0,0675
3	Selisih Belanja Pegawai (Proxy TPG/TKG)	0,0149
4	Nominal Belanja Pegawai Total (Rp)	0,0032
5	Nominal Belanja Pegawai Murni (Rp)	0,0028
6	Tahun Anggaran (2025/2026)	0,0021
7	Nominal Belanja Daerah (Rp)	0,0019

Sumber: penelitian ini

Distribusi Hasil Klasifikasi dan Implikasi Kebijakan

Setelah model dilatih dan divalidasi, model SVM digunakan untuk mengklasifikasikan seluruh 1.092 data daerah. Hasil menunjukkan bahwa pada TA 2025, 105 daerah (19,2%) diprediksi Efisien dan 441 daerah (80,8%) Tidak Efisien. Pada TA 2026, jumlah daerah Efisien turun menjadi 69 (12,6%), sementara daerah Tidak Efisien meningkat menjadi 477 (87,4%). Tren peningkatan proporsi daerah Tidak Efisien dari 80,8% menjadi 87,4% ini memberikan landasan ilmiah bagi Kementerian Keuangan dan Kementerian Dalam Negeri untuk menerapkan langkah-langkah restrukturisasi belanja pegawai daerah, termasuk rasionalisasi struktur organisasi dan penerapan ketentuan UU HKPD secara konsisten [2].

Perbandingan dengan Penelitian Terdahulu

Untuk memposisikan capaian penelitian ini terhadap studi-studi sejenis yang telah dibahas pada tinjauan literatur, Tabel 11 merangkum perbandingan akurasi model SVM pada berbagai domain data pemerintahan dan publik.

Tabel 11. Perbandingan Akurasi Model SVM dengan Penelitian Terdahulu

Peneliti	Objek Penelitian	Algoritma	Akurasi
Yazid et al. [3]	Rekomendasi Bansos APBD Kota Pasuruan	SVM Linear	95%
Puspandoyo et al. [5]	Kualitas Laporan Keuangan Kementerian	SVM	96,1%
Yusmita et al. [10]	Sentimen Kebijakan Efisiensi Anggaran	SVM LinearSVC	94%

Prasetyo & Maulana [7]	Potensi Tsunami di Indonesia	SVM Polynomial	97%
Ramadhani et al. [8]	Diagnosis Kesehatan Berbasis Web	SVM Linear	99%
Rizki & Darip [15]	Anomali Aktivitas Sistem Keuangan	SVM RBF	86%
Penelitian ini	Efisiensi Belanja Pegawai Daerah	SVM RBF (C=100, $\gamma=1$)	98,63%

Sumber: penelitian ini

Akurasi 98,63% yang dicapai penelitian ini berada di atas rata-rata capaian studi-studi sejenis pada domain keuangan dan layanan publik, sekaligus mengonfirmasi bahwa kombinasi kernel RBF dengan optimasi parameter melalui Grid Search 5-Fold Cross Validation mampu menangkap pola non-linear pada data belanja pegawai daerah secara lebih presisi dibandingkan konfigurasi SVM yang lebih sederhana [3][5][7]. Perbandingan ini juga menegaskan bahwa performa tinggi bukan sekadar hasil dari satu dataset tunggal, melainkan konsisten dengan tren keunggulan SVM pada berbagai karakteristik data keuangan dan sosial pemerintahan yang telah dilaporkan pada literatur terdahulu [8][10][15].

KESIMPULAN

Penelitian ini berhasil membangun model klasifikasi berbasis Support Vector Machine untuk mengelompokkan tingkat efisiensi belanja pegawai pemerintah daerah di Indonesia. Model SVM dengan konfigurasi kernel RBF, C=100, dan gamma=1 yang diperoleh melalui Grid Search 5-Fold Cross Validation terbukti mampu mengklasifikasikan tingkat efisiensi belanja pegawai dengan performa yang sangat tinggi dan stabil, mencapai akurasi 98,63%, presisi 98,62%, recall 98,63%, dan F1-Score 98,62% pada data pengujian. Berdasarkan capaian ini, hipotesis H1 diterima: terdapat kemampuan diskriminasi yang signifikan dari model SVM dalam mengklasifikasikan efisiensi belanja pegawai daerah.

Analisis feature importance mengidentifikasi bahwa Rasio Belanja Pegawai Murni merupakan faktor paling dominan dalam menentukan klasifikasi tingkat efisiensi, dengan rata-rata penurunan F1-Score sebesar 0,1977. Model memprediksi bahwa 87,4% daerah pada TA 2026 terklasifikasi Tidak Efisien, mencerminkan tren peningkatan beban belanja pegawai dalam struktur APBD yang semakin mengkhawatirkan. Temuan ini memberikan landasan ilmiah yang kuat untuk penetapan prioritas intervensi kebijakan restrukturisasi belanja pegawai secara berbasis data [2][11].

Hasil evaluasi lebih lanjut melalui confusion matrix menunjukkan tingkat kesalahan klasifikasi yang sangat rendah, yakni hanya 3 dari 219 data pengujian (1,37%) yang salah diklasifikasikan, dengan performa terbaik pada kelas Tidak Efisien (F1-Score 99,19%) [7]. Temuan ini menegaskan bahwa model SVM yang dibangun tidak hanya akurat secara agregat, tetapi juga konsisten dalam membedakan kedua kelas efisiensi meskipun distribusi data antarkelas tidak seimbang.

Penelitian selanjutnya disarankan untuk mengintegrasikan data realisasi APBD serta menambahkan variabel kontekstual seperti IPM, PDRB per kapita, dan luas wilayah. Selain itu, komparasi dengan algoritma lain seperti Random Forest, XGBoost, atau LightGBM perlu dilakukan untuk mendapatkan metode klasifikasi yang paling optimal. Pengembangan dashboard monitoring

berbasis web yang terintegrasi dengan SIPD juga direkomendasikan agar klasifikasi efisiensi belanja pegawai dapat dilakukan secara otomatis dan real-time [1][11].

REFERENSI

- R. Sholeha, A. Zahrah, B. Martapuse, and P. Pujianto, "Prediksi Jumlah Belanja Pegawai Provinsi Sumatra Selatan," *Pros. Semin. Nas. Teknol. Inf. dan Bisnis*, pp. 377–385, 2025, doi: 10.47701/3thr6h28.
- J. Rybak, H. Battey, and W. X. Zhou, "On Inference for the Support Vector Machine," *J. Mach. Learn. Res.*, vol. 26, pp. 1–54, 2025.
- M. Yazid, D. A. Sani, and N. M. Anggadimas, "Sistem Rekomendasi Penerima Bantuan Sosial APBD Menggunakan Metode Support Vector Machine," *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 6, no. 1, pp. 54–61, 2024, doi: 10.28926/ilkomnika.v6i1.619.
- E. R. S. Moningkey, D. D. Harisonak, and K. Santa, "Support Vector Machine Algorithm for Classifying Public Satisfaction Index," *Acad. Open*, vol. 10, no. 2, 2025, doi: 10.21070/acopen.10.2025.12697.
- T. Puspadoyo et al., "Prediksi Kualitas/Laporan Keuangan Kementerian Negara /Lembaga Menggunakan," vol. 3, pp. 15–36, 2022.
- Rofiq, "Penegenalan Dasar Analisis Data Dengan Python Di Goggle Colab," 2024.
- H. Prasetyo and A. E. Maulana, "Penerapan Algoritma Support Vector Machine untuk Melakukan Analisis Potensi Tsunami di Indonesia," *Multinetics*, vol. 10, no. 2, pp. 174–182, 2025, doi: 10.32722/multinetics.v10i2.7245.
- M. A. Ramadhani et al., "Implementasi Algoritma Support Vector Machine (SVM) Untuk Diagnosis Kesehatan Manusia Berbasis Web Application," *J. Ners*, vol. 9, no. 1, pp. 896–902, 2025.
- B. Santoso, *Support Vector Machine (Algoritma dan Aplikasinya)*. Karya Buku dan Jurnal Indonesia, 2025.
- Yusmita, Farid Wajidi, and Muh.Rafli Rasyid, "Comparison of SVM and Naive Bayes in Public Sentiment Analysis on Budget Efficiency," *J. Sist. Cerdas*, vol. 8, no. 3, pp. 332–342, 2025, doi: 10.37396/jsc.v8i3.576.
- I. Mahendro and D. Abimanto, *Algoritma Support Vector Machine untuk Menganalisa Kepuasan Mahasiswa terhadap Pembelajaran E-Learning*. CV Pustaka Stimart AMNI, 2023.
- U. B. Muhamad, Rohmat Nur, "Analisi Sentimen Public Terhadap Program Bantuan Subsidi Upah (BSU) Di Twitter Menggunakan Algoritma Svm Tugas Akhir Oleh: Program Studi Teknik Informatika Universitas Budi Luhur Jakarta," vol. 8, pp. 2–9, 2025.
- D. Purnamasari et al., *Pengantar Metode Analisis Sentimen*. 2023.
- W. L. Azzahra and E. Mailoa, "Analisis Sentimen terhadap RSUD Salatiga Menggunakan SVM dan TF-IDF," *J. Indones. Manaj. Inform. dan Komun.*, vol. 6, no. 1, pp. 478–489, 2025, doi: 10.35870/jimik.v6i1.1208.

W. Rizki and M. Darip, “Penggunaan Algoritma Support Vector Machine Untuk Mendeteksi Anomali Aktivitas Pengguna Pada Sistem Informasi Keuangan Pt. Digidokat Indonesia,” JATI (Jurnal Mhs. Tek. Inform., vol. 9, no. 3, pp. 4538–4546, 2025, doi: 10.36040/jati.v9i3.13385.