

Implementation of the Support Vector Machine Algorithm for Predicting Voter Participation Based on Village Development Index Data in the Tapal Kuda Region

[Penerapan Algoritma Support Vector Machine dalam Memprediksi Tingkat Partisipasi Pemilu Berdasarkan Data Indeks Desa Membangun pada Wilayah Tapal Kuda]

M. Bahrul Ulum¹⁾, Arif Senja Fitriani²⁾, Hindarto³⁾, Uce Indahyanti⁴⁾

^{1),(2),(3),(4)} Program Studi Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia

*Email Penulis Korespondensi: asfjim@umsida.ac.id

Abstract. *The level of voter participation reflects not only the strength of democracy but also the social awareness of the community. In East Java's Tapal Kuda region, this participation varies significantly between villages, influenced by diverse development conditions. This study explores the use of the Village Development Index (Indeks Desa Membangun/IDM) to predict voter participation through the Support Vector Machine (SVM) algorithm. Using the 2024 IDM dataset consisting of thousands of village-level records, several indicators were selected to represent social, economic, and infrastructural aspects. The SVM model produced an accuracy of 76%, with precision and recall values of 0.75 and 0.76 respectively. These results indicate that the model is more responsive in detecting high participation patterns, suggesting that village development and access to basic services contribute significantly to civic engagement. Therefore, the application of SVM on IDM data can be utilized as a data-driven approach to support strategies aimed at strengthening voter participation across rural areas.*

Keywords - oter Participation; Village Development Index; Support Vector Machine; Machine Learning; Tapal Kuda.

Abstrak *Tingkat partisipasi pemilih mencerminkan tidak hanya kualitas demokrasi, tetapi juga kesadaran sosial masyarakat. Di wilayah Tapal Kuda, Jawa Timur, variasi partisipasi antar desa menunjukkan adanya pengaruh dari kondisi pembangunan yang beragam. Penelitian ini mengkaji penerapan Indeks Desa Membangun (IDM) untuk memprediksi tingkat partisipasi pemilih menggunakan algoritma Support Vector Machine (SVM). Data penelitian berasal dari publikasi IDM tahun 2024 yang memuat ribuan data desa dengan indikator sosial, ekonomi, dan infrastruktur. Hasil analisis menunjukkan bahwa model SVM mencapai akurasi sebesar 76%, dengan nilai precision 0,75 dan recall 0,76. Model terbukti lebih efektif dalam mengenali pola partisipasi tinggi, menandakan bahwa kemajuan desa dan akses terhadap layanan dasar berperan penting dalam meningkatkan partisipasi masyarakat. Temuan ini menegaskan bahwa pendekatan SVM berbasis data IDM dapat menjadi landasan bagi perumusan strategi kebijakan untuk mendorong keterlibatan politik masyarakat, khususnya di wilayah pedesaan.*

Kata Kunci - Partisipasi Pemilih; Indeks Desa Membangun; Support Vector Machine; Machine Learning; Tapal Kuda.

I. PENDAHULUAN

Demokrasi adalah sistem pemerintahan yang menempatkan rakyat sebagai subjek utama dalam pengambilan keputusan politik. Dalam sistem ini, partisipasi warga negara bukan hanya menjadi hak, tetapi juga tanggung jawab kolektif untuk menentukan arah pembangunan suatu negara. Salah satu bentuk partisipasi yang paling konkret dan krusial adalah melalui pemilihan umum, yang sekaligus menjadi indikator kualitas demokrasi di suatu negara [1].

Di Indonesia, meskipun tingkat partisipasi dalam pemilu terbilang tinggi, masih terdapat sejumlah tantangan yang perlu diperhatikan. Berdasarkan data Komisi Pemilihan Umum (KPU), partisipasi pemilih dalam Pemilu Presiden 2019 mencapai 80,76%. Namun, angka ini juga menunjukkan bahwa sekitar 19,24% pemilih tidak menggunakan hak suaranya [2]. Berbagai faktor seperti rendahnya literasi politik, keterbatasan akses informasi, ketidakpercayaan terhadap proses demokrasi, dan hambatan geografis menjadi penyebab utama dari kondisi tersebut [3]. Oleh karena itu, diperlukan analisis yang lebih mendalam untuk memahami faktor-faktor yang memengaruhi partisipasi pemilih, khususnya di wilayah dengan kondisi sosial dan geografis yang kompleks.

Salah satu wilayah yang menarik untuk dikaji adalah kawasan Tapal Kuda di Jawa Timur, yang terdiri dari tujuh kabupaten: Banyuwangi, Bondowoso, Jember, Lumajang, Pasuruan, Situbondo, dan Probolinggo. Kawasan ini dikenal dengan keragaman budaya, kondisi sosial ekonomi yang belum merata, serta keterbatasan infrastruktur di beberapa daerah. Banyak desa di wilayah ini masih berada dalam kategori berkembang bahkan tertinggal, yang sangat mungkin turut memengaruhi partisipasi politik masyarakatnya.

Untuk melihat hubungan antara kondisi pembangunan desa dan partisipasi masyarakat dalam pemilu, salah satu instrumen yang dapat digunakan adalah Indeks Desa Membangun (IDM) yang dikembangkan oleh Kementerian Desa.

IDM mencakup tiga aspek utama: ketahanan sosial, ketahanan ekonomi, dan ketahanan lingkungan, dan digunakan untuk mengklasifikasikan desa dari kategori sangat tertinggal hingga mandiri [4]. Dengan memanfaatkan data IDM, peneliti dapat melakukan pendekatan berbasis data untuk menilai apakah terdapat korelasi antara tingkat pembangunan dan partisipasi pemilu.

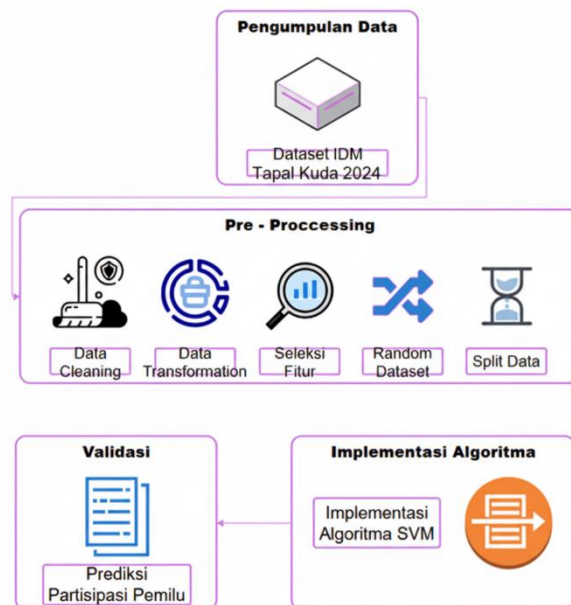
Dalam konteks pengolahan data dan pembuatan model prediksi, algoritma Support Vector Machine (SVM) dipilih karena kemampuannya dalam mengklasifikasikan data baik linear maupun non-linear, serta dalam membentuk model prediksi yang akurat. Pada penelitian terdahulu, SVM telah terbukti efektif dalam memprediksi partisipasi pemilu melalui berbagai pendekatan, seperti analisis hubungan antara tingkat pendidikan dan partisipasi politik, serta analisis sentimen publik di media sosial terhadap isu-isu pemilu [5]. Secara umum, hasil penelitian tersebut menunjukkan bahwa SVM mampu menangani data yang kompleks dan menghasilkan klasifikasi yang memadai. Sementara itu, pendekatan lain seperti algoritma K-Means lebih berfokus pada pengelompokan data dan belum sepenuhnya efektif untuk prediksi terhadap target yang spesifik [6].

Namun, sebagian besar penelitian terdahulu masih terbatas dalam penggunaan variabel dan wilayah kajian [6]. Belum banyak yang mengintegrasikan data IDM sebagai variabel utama untuk memprediksi partisipasi politik di wilayah pedesaan, khususnya Tapal Kuda. Di sinilah letak kebaruan dan kontribusi penelitian ini. Dengan menggabungkan data IDM dan algoritma SVM, diharapkan hasil penelitian ini mampu menyajikan gambaran prediksi yang lebih akurat dan kontekstual, serta dapat menjadi referensi dalam menyusun strategi peningkatan partisipasi pemilih di wilayah yang masih berkembang atau tertinggal.

II. METODE

Tahapan Penelitian

Tahapan penelitian merupakan langkah-langkah sistematis yang dilakukan mulai dari perencanaan awal hingga evaluasi akhir guna mencapai tujuan penelitian. Pada penelitian ini, proses dimulai dari studi literatur, dilanjutkan dengan pengumpulan data demografi dari tujuh kabupaten di wilayah Tapal Kuda, yaitu Banyuwangi, Bondowoso, Jember, Lumajang, Pasuruan, Situbondo, dan Probolinggo. Data yang digunakan berjumlah lebih dari 1000 instance dengan 48 atribut. Selanjutnya dilakukan tahap pre-processing yang mencakup data cleaning, transformasi data, seleksi fitur, randomisasi, dan split data [7]. Data yang telah diproses kemudian dianalisis menggunakan algoritma SVM, dengan klasifikasi target berupa sentimen partisipasi masyarakat terhadap presiden terpilih. Rangkaian tahapan penelitian ditunjukkan pada Gambar 1 berikut:



Gambar 1. Tahapan Penelitian.

Rangkaian tahapan penelitian ditunjukkan pada Gambar 1 berikut, yang menjelaskan alur penelitian dari tahap awal hingga hasil :

1. Pengumpulan Data

Pengumpulan data dilakukan pada tahap awal penelitian dengan fokus pada wilayah Tapal Kuda di Provinsi Jawa Timur, yang mencakup tujuh kabupaten: Banyuwangi, Bondowoso, Jember, Lumajang, Pasuruan, Situbondo, dan Probolinggo. Wilayah ini dipilih karena dinilai strategis untuk dianalisis dalam konteks prediksi partisipasi pemilu

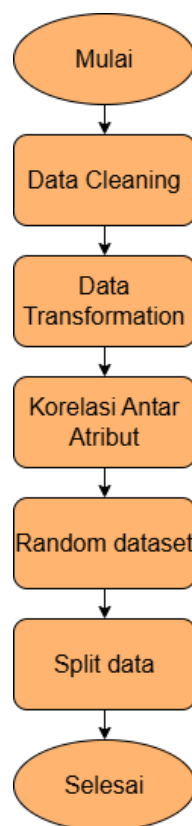
berdasarkan karakteristik geografis dan demografisnya [8]. Data yang digunakan merupakan data publikasi tahun 2024 yang terdiri dari lebih dari 1000 entri, mencakup 48 atribut prediktor dan 1 atribut target berupa klasifikasi sentimen partisipasi pemilih terhadap presiden terpilih. Seluruh data kemudian diolah dalam tahap preprocessing untuk mendukung proses klasifikasi menggunakan algoritma SVM. Rincian atribut yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Atribut Data

No Atribut	Kelompok Atribut	Atribut
X1 – X4	Pendidikan	Akses terhadap PAUD/ TK/ Sederajat (X1), Akses terhadap SD/ MI/ Sederajat (X2), Akses terhadap SMP/ MTs/ Sederajat (X3), Akses terhadap SMA/ SMK/ MA/ MAK/ Sederajat (X4)
X5 – X11	Kesehatan	Layanan Sarana Kesehatan (X5), Fasilitas Kesehatan Poskesdes/ Polindes (X6), Aktivitas Posyandu (X7), Layanan Dokter (X8), Layanan Bidan (X9), Layanan Tenaga Kesehatan Lainnya (X10), Jaminan Kesehatan Nasional (X11)
X12 - X13	Utilitas dasar	Air Minum (X12), Persentase Rumah Tidak Layak Huni (X13)
X14 – X18	Aktivitas	Kearifan Sosial/ Budaya (X14), Frekuensi Gotong Royong (X15), Kegiatan Olahraga (X16), Mitigasi dan Penanganan Konflik Sosial (X17), Satkamling (X18)
X19 - X21	Fasilitas masyarakat	Taman Bacaan Masyarakat/ Perpustakaan Desa (X19), Fasilitas Olahraga (X20), Keberadaan Ruang Publik Terbuka (X21)
X22 – X25	Produksi desa	Keragaman Aktivitas Ekonomi (X22), Produk Unggulan Desa (X23), Ekonomi Kreatif (X24), Kerjasama Desa (X25)
X26 – X33	Pendukung ekonomi	Akses Terhadap Pendidikan Non-formal/ Pusat Keterampilan/ Kursus (X26), Pasar Rakyat (X27), Toko/ Pertokoan (X28), Kedai/ Rumah Makan (X29), Penginapan (X30), Layanan Pos dan/ Logistik (X31), Lembaga Ekonomi (X32), Layanan Keuangan (X33)
X34 – X37	Pengelolaan lingkungan	Kearifan Lingkungan (X34), Sistem Pengelolaan Sampah (X35), Tingkat Pencemaran Lingkungan (X36), Sistem Pembuangan Air Limbah Domestik (Rumah Tangga) (X37)
X38	Penanggulangan bencana	Penanggulangan Bencana (X38)
X39 – X40	Kondisi akses jalan	Kondisi Jalan di desa (X39), Kondisi Penerangan Jalan Utama Desa (X40)

X41 – X43	Kemudahan akses	Keberadaan Angkutan Perdesaan/ Angkutan Lokal/ Sejenis (X41), Akses Listrik (X42), Layanan Telekomunikasi (X43)
X44 – X46	Kelembagaan dan pelayanan desa	Pelaksanaan Pelayanan dan Administrasi Desa (X44), Pemanfaatan Teknologi dalam Pelayanan Desa (SPBE) (X45), Musyawarah Desa (X46)
X47 - X48	Tata kelola keuangan desa	Pendapatan Asli Desa (PADes) dan Dana Desa (X47), Jumlah Kepemilikan dan Produktivitas Aset Desa (X48)
Y	Pemilu	Tingkat Partisipasi Masyarakat

2. Pre-Processing



Gambar 2. Flowchart Pre-Processing

Pre-processing data merupakan tahap awal dalam implementasi algoritma SVM pada data Indeks Desa Membangun, yang meliputi penggabungan seluruh data, modifikasi, pembersihan, dan penyederhanaan data untuk mendukung prediksi partisipasi pemilu di wilayah Tapal Kuda.

a Data Cleaning

Sebelum digabung, dilakukan normalisasi teks pada kolom wilayah agar penulisan antar-dataset konsisten [9]. Normalisasi meliputi perubahan huruf menjadi lowercase, penghapusan kata administratif seperti “desa”, “kecamatan”, dan “kabupaten”, serta pembersihan spasi di awal maupun akhir string. Dataset kemudian digabung menggunakan `pd.merge()` dengan Kabupaten, Kecamatan, dan Desa sebagai key penghubung. Tipe join yang digunakan adalah inner join, sehingga hanya baris yang memiliki kecocokan pada ketiga kolom tersebut yang dipertahankan. Tahap ini juga meliputi penanganan duplikasi, pengisian atau penghapusan missing values, serta perbaikan data yang tidak konsisten. Selain itu, atribut yang bersifat administratif seperti nama wilayah, kode wilayah, dan tahun dihapus karena tidak relevan untuk analisis.

b Data Transformation

Seluruh variabel kemudian ditransformasikan agar sesuai dengan kebutuhan analisis. Proses transformasi ini meliputi konversi data kategorikal menjadi numerik, misalnya persentase partisipasi yang sebelumnya disajikan dalam bentuk teks diubah ke dalam tipe numerik agar dapat diolah oleh algoritma [10]. Selain itu, dilakukan pula normalisasi atau standarisasi data untuk menyamakan skala antarvariabel sehingga tidak menimbulkan bias dalam proses pemodelan.

c Seleksi Fitur

Dalam implementasi algoritma SVM pada data IDM untuk wilayah Tapal Kuda, seleksi fitur menjadi langkah penting untuk menyederhanakan struktur data dan meningkatkan akurasi model. Dari total 48 atribut prediktor yang tersedia, dilakukan peninjauan terhadap relevansi masing-masing atribut terhadap target prediksi, yaitu partisipasi pemilu [11]. Atribut-atribut yang dinilai memiliki pengaruh rendah atau redundan dieliminasi dari dataset. Hasil dari proses ini menyisakan 27 atribut yang dianggap paling representatif untuk mendukung proses klasifikasi. Rincian atribut yang digunakan dalam penelitian ini disajikan pada Tabel 2.

Tabel 2. Tabel Atribut Hasil Seleksi Fitur

No Atribut	Kelompok Atribut	Atribut
X1 – X4	Pendidikan	Akses terhadap PAUD/ TK/ Sederajat (X1), Akses terhadap SD/ MI/ Sederajat (X2), Akses terhadap SMP/ MTs/ Sederajat (X3), Akses terhadap SMA/ SMK/ MA/ MAK/ Sederajat (X4)
X5 – X11	Kesehatan	Layanan Sarana Kesehatan (X5), Fasilitas Kesehatan Poskesdes/ Polindes (X6), Aktivitas Posyandu (X7), Layanan Dokter (X8), Layanan Bidan (X9), Layanan Tenaga Kesehatan Lainnya (X10), Jaminan Kesehatan Nasional (X11)
X12 – X16	Aktivitas	Kearifan Sosial/ Budaya (X12), Frekuensi Gotong Royong (X13), Kegiatan Olahraga (X14), Mitigasi dan Penanganan Konflik Sosial (X15), Satkamling (X16)
X17 – X19	Fasilitas masyarakat	Taman Bacaan Masyarakat/ Perpustakaan Desa (X17), Fasilitas Olahraga (X18), Keberadaan Ruang Publik Terbuka (X19)
X20 – X21	Kondisi akses jalan	Kondisi Jalan di desa (X20), Kondisi Penerangan Jalan Utama Desa (X21)
X22 – X24	Kemudahan akses	Keberadaan Angkutan Perdesaan/ Angkutan Lokal/ Sejenis (X22), Akses Listrik (X23), Layanan Telekomunikasi (X24)
X25 – X27	Kelembagaan dan pelayanan desa	Pelaksanaan Pelayanan dan Administrasi Desa (X25), Pemanfaatan Teknologi dalam Pelayanan Desa (SPBE) (X26), Musyawarah Desa (X27)
Y	Pemilu	Tingkat Partisipasi Masyarakat

d Korelasi Antar Atribut

Tahap berikutnya adalah melakukan analisis korelasi antar fitur terhadap 27 atribut hasil seleksi pertama. Analisis ini bertujuan untuk mengidentifikasi hubungan antarvariabel serta mengurangi redundansi data. Atribut yang memiliki korelasi negatif, terlalu lemah, atau tidak signifikan terhadap variabel target dianggap tidak memberikan kontribusi berarti dalam proses prediksi, sehingga perlu dieliminasi. Dengan menghapus atribut-atribut tersebut, dataset menjadi lebih ringkas, relevan, dan bebas dari informasi ganda yang berpotensi menurunkan performa model. Hasil dari tahap ini menyisakan atribut inti yang dinilai paling representatif untuk mendukung proses klasifikasi menggunakan algoritma SVM.

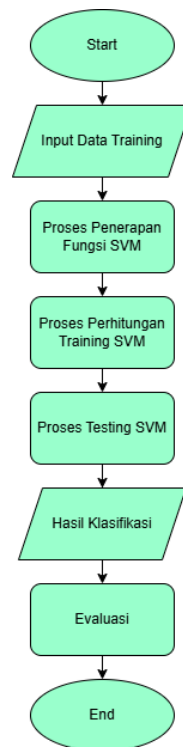
e Random dataset

Sebelum proses pelatihan model dilakukan, data diacak terlebih dahulu guna memastikan distribusi atribut yang lebih merata dan menghindari bias urutan pada dataset.

f Split data

Setelah melalui tahap preprocessing dan seleksi fitur, data dibagi menjadi dua bagian, yaitu 70% untuk data pelatihan (training) dan 30% untuk data pengujian (testing), agar model dapat diuji secara objektif terhadap data yang belum dikenali sebelumnya.

3. Process



Gambar 3. Flowchart Tahapan Pemrosesan Data

Flowchart di atas menggambarkan tahapan pemrosesan data dalam implementasi algoritma SVM pada data IDM untuk memprediksi partisipasi pemilu di wilayah Tapal Kuda. Berikut penjelasan dari flowchart diatas :

a Start

Proses dimulai dengan menyiapkan data yang akan digunakan sebagai input.

b Input Data Training

Data yang berasal dari Indeks Desa Membangun dimasukkan sebagai data training. Data ini telah melalui tahap pre-processing seperti penggabungan, pembersihan, normalisasi, dan transformasi agar sesuai dengan format yang dibutuhkan oleh algoritma SVM.

c Proses Penerapan Fungsi SVM

Fungsi SVM mulai dijalankan untuk menemukan hyperplane terbaik yang dapat memisahkan data berdasarkan label partisipasi pemilu [12].

d Proses Perhitungan Training SVM

Pada tahap ini, sistem melakukan proses pembelajaran (training) terhadap data yang telah dilabeli. Algoritma SVM akan menentukan parameter optimal yang memaksimalkan margin antar kelas [13].

e Proses Testing SVM

Setelah proses training selesai, dilakukan pengujian (testing) terhadap data uji untuk mengukur kinerja model dalam mengklasifikasikan data yang belum dikenali sebelumnya.

f Hasil Klasifikasi

Sistem menghasilkan hasil klasifikasi berupa prediksi tingkat partisipasi pemilu berdasarkan input karakteristik desa dari data IDM.

g End

Proses pemrosesan data selesai dan model siap digunakan untuk prediksi pada data-data lain.

4. Output

Tahap ini menghasilkan keluaran berupa prediksi tingkat partisipasi masyarakat dalam pemilu yang diklasifikasikan ke dalam dua kategori, yaitu “Tinggi” dan “Rendah”. Model SVM menghasilkan klasifikasi berdasarkan hyperplane terbaik yang telah dibentuk selama proses pelatihan untuk memisahkan data sesuai dengan karakteristiknya.

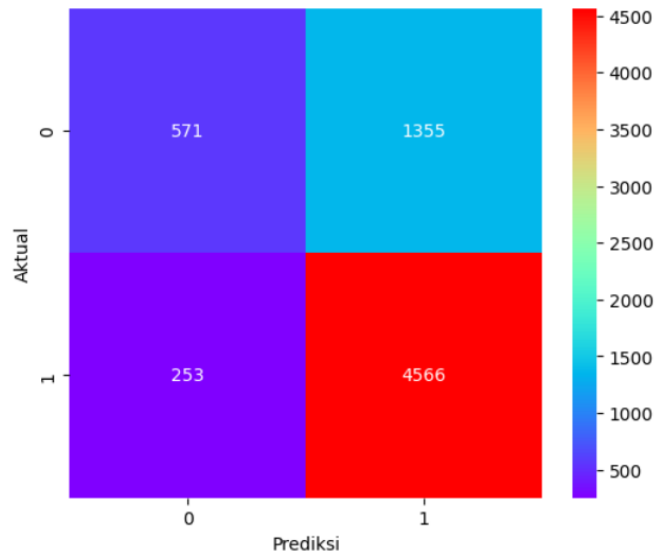
5. Analisis / Evaluasi

Tahap evaluasi dilakukan untuk menilai kinerja model SVM yang telah dibangun. Pengujian dilakukan menggunakan beberapa metrik evaluasi, yaitu akurasi, presisi, recall, dan F1-score, guna mengukur sejauh mana

model mampu mengklasifikasikan partisipasi masyarakat dengan benar. Hasil dari evaluasi ini memberikan gambaran mengenai efektivitas algoritma SVM dalam memprediksi tingkat partisipasi pemilu berdasarkan data IDM.

III. HASIL DAN PEMBAHASAN

Akurasi dan Presisi



Gambar 2. Confusion Matrix

Berdasarkan confusion matrix pada gambar di atas, model menunjukkan performa klasifikasi yang cukup baik dengan jumlah prediksi benar yang tinggi pada kelas positif (4566 data teridentifikasi dengan benar). Nilai True Positive (TP) sebesar 4566 dan True Negative (TN) sebesar 571 menunjukkan bahwa model mampu mengenali sebagian besar data dengan tepat. Namun, masih terdapat kesalahan klasifikasi berupa False Positive (FP) sebanyak 1355 dan False Negative (FN) sebanyak 253. Secara umum, hasil ini mencerminkan tingkat akurasi dan presisi yang cukup tinggi, di mana model lebih baik dalam mendeteksi kelas positif dibandingkan kelas negatif [14]. Hal ini juga mengindikasikan bahwa model cenderung lebih sensitif terhadap kelas positif, sehingga cocok digunakan ketika prioritasnya adalah mengurangi kesalahan pada prediksi positif.

Classification Report

Tabel 4. Classification Report SVM

Kelas	Precision	Recall	F1-Score	Support
RENDAH	0.69	0.30	0.42	1,926
TINGGI	0.77	0.95	0.85	4,819
Accuracy			0.76	6,745
Macro Avg	0.73	0.62	0.63	6,745
Weighted Avg	0.75	0.76	0.73	6,745

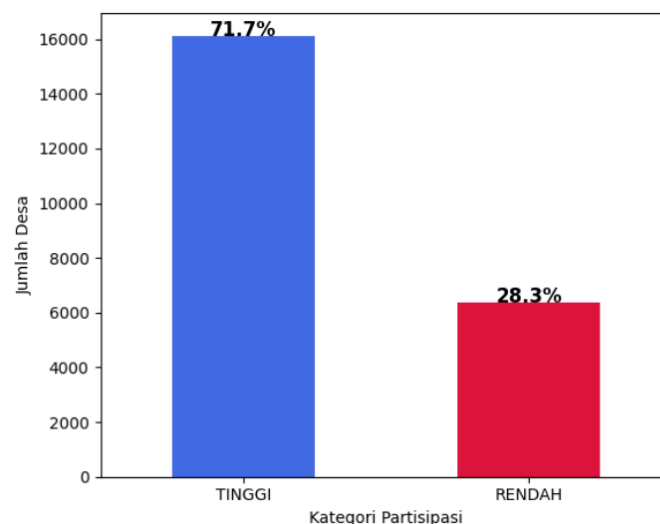
Berdasarkan hasil evaluasi pada tabel classification report, model memperoleh nilai akurasi sebesar 0,76, yang berarti 76% dari seluruh data uji berhasil diprediksi dengan benar. Pada kelas Tinggi, nilai precision sebesar 0,77 dan recall sebesar 0,95 menunjukkan bahwa model sangat baik dalam mengenali dan mengklasifikasikan data pada kelas ini. Sementara itu, pada kelas Rendah, nilai precision sebesar 0,69 dan recall sebesar 0,30 menandakan bahwa masih terdapat cukup banyak data dari kelas Rendah yang salah terprediksi sebagai Tinggi. Nilai F1-score sebesar 0,85 pada kelas Tinggi dan 0,42 pada kelas Rendah memperkuat temuan ini, di mana model lebih optimal dalam mendeteksi kelas mayoritas. Secara keseluruhan, nilai weighted average precision 0,75 dan recall 0,76 menunjukkan bahwa performa model tergolong baik dan stabil.

Tabel 5. Classification Report Decision Tree

Kelas	Precision	Recall	F1-Score	Support
RENDAH	0.36	0.64	0.46	1926
TINGGI	0.79	0.54	0.64	4819
Accuracy			0.57	6745
Macro Avg	0.58	0.59	0.55	6745
Weighted Avg	0.67	0.57	0.59	6745

Sebagai pembandingan terhadap model SVM, hasil evaluasi pada Decision Tree menunjukkan bahwa performanya masih kurang stabil dan tidak merata antar kelas. Pada kelas rendah, model menghasilkan recall yang cukup tinggi (0,64), namun precision yang rendah (0,36) menandakan banyak prediksi yang keliru. Sementara itu, pada kelas TINGGI, meskipun precision mencapai 0,79, nilai recall hanya 0,54, sehingga sebagian besar data kelas mayoritas tidak teridentifikasi dengan baik. Ketidakseimbangan performa ini berkontribusi terhadap accuracy keseluruhan yang hanya mencapai 0,57. Jika dibandingkan dengan SVM, Decision Tree terlihat kurang mampu menjaga konsistensi dan ketepatan prediksi. SVM memiliki nilai accuracy, F1-score, dan weighted average yang lebih tinggi, menandakan kualitas generalisasi yang lebih baik. Dengan demikian, hasil ini menunjukkan bahwa SVM lebih unggul dan lebih dapat diandalkan dibandingkan Decision Tree dalam melakukan klasifikasi pada dataset ini.

Presentase Tingkat Partisipasi

**Gambar 2.** Presentase Tingkat Partisipasi

Gambar 3 menunjukkan distribusi kategori partisipasi desa yang digunakan dalam penelitian ini. Berdasarkan grafik tersebut, terlihat bahwa mayoritas data berada pada kategori partisipasi tinggi sebesar 71,7%, sedangkan kategori partisipasi rendah hanya mencapai 28,3%. Ketidakseimbangan data ini menunjukkan bahwa data bersifat tidak seimbang (imbalanced), di mana kelas Tinggi memiliki jumlah data jauh lebih banyak dibandingkan kelas Rendah. Kondisi tersebut dapat memengaruhi performa model klasifikasi karena algoritma cenderung lebih mudah mengenali pola dari kelas mayoritas [15]. Meskipun teknik penyeimbangan data seperti SMOTE telah diterapkan untuk mengatasi perbedaan jumlah data antar kelas, hasil pengujian menunjukkan bahwa perbedaan performa antar kelas masih tetap muncul. Hal ini mengindikasikan bahwa karakteristik data pada kelas Rendah memang lebih kompleks atau memiliki pola yang sulit dikenali, bukan semata karena ketidakseimbangan data. Dengan demikian, performa model yang diperoleh tetap dapat dikategorikan baik dan representatif terhadap kondisi data sebenarnya.

VII. SIMPULAN

Berdasarkan hasil pengujian dan evaluasi model klasifikasi menggunakan metode Support Vector Machine (SVM), dapat disimpulkan bahwa model memiliki kinerja yang baik dan stabil dalam mengklasifikasikan tingkat partisipasi desa. Berdasarkan Confusion Matrix, model berhasil mengidentifikasi sebagian besar data dengan benar, dengan nilai True Positive (TP) sebesar 4.566 dan True Negative (TN) sebesar 571. Meskipun masih terdapat

kesalahan klasifikasi pada False Positive (FP) sebanyak 1.355 dan False Negative (FN) sebanyak 253, hasil ini menunjukkan bahwa model memiliki kemampuan yang cukup baik dalam mengenali pola data terutama pada kelas partisipasi tinggi.

Hasil Classification Report memperkuat temuan tersebut, di mana model memperoleh nilai akurasi sebesar 0,76, dengan precision sebesar 0,77 dan recall sebesar 0,95 pada kelas partisipasi tinggi. Nilai tersebut menunjukkan bahwa model sangat efektif dalam mengklasifikasikan kelas mayoritas, sementara performa pada kelas partisipasi rendah masih lebih rendah dengan recall sebesar 0,30. Perbedaan ini berkaitan dengan distribusi data yang tidak seimbang (imbalanced), di mana kelas partisipasi tinggi mendominasi sebanyak 71,7% dari total data, sedangkan kelas partisipasi rendah hanya sebesar 28,3%.

Meskipun teknik SMOTE (Synthetic Minority Over-sampling Technique) telah diterapkan untuk mengatasi ketidakseimbangan data, variasi performa antar kelas masih terjadi. Hal ini menunjukkan bahwa karakteristik data pada kelas partisipasi rendah memiliki pola yang lebih kompleks dan sulit dikenali oleh model. Secara keseluruhan, hasil pengujian menunjukkan bahwa metode SVM mampu memberikan hasil prediksi yang representatif terhadap kondisi data aktual, dengan tingkat akurasi yang dapat diterima. Dengan demikian, model yang dikembangkan dapat dinyatakan efektif dalam mengklasifikasikan tingkat partisipasi desa serta memiliki potensi untuk digunakan dalam analisis atau pengambilan keputusan pada konteks serupa di masa mendatang.

UCAPAN TERIMA KASIH

Penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah memberikan dukungan, arahan, serta motivasi selama proses penyusunan penelitian ini. Apresiasi khusus diberikan kepada Pemerintah Kabupaten Sidoarjo atas ketersediaan data dan dukungannya dalam penelitian mengenai partisipasi pemilu. Ucapan terima kasih juga disampaikan kepada dosen pembimbing yang telah memberikan bimbingan dan masukan berharga, serta seluruh dosen Program Studi Informatika Universitas Muhammadiyah Sidoarjo atas ilmu dan inspirasinya. Tak lupa, penulis berterima kasih kepada keluarga dan teman-teman atas doa, dorongan, dan semangat yang tiada henti. Semoga penelitian ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan serta menjadi kontribusi nyata dalam upaya meningkatkan partisipasi politik masyarakat.

REFERENSI

- [1] K. Kunci, "The Indonesian Journal of Computer Science," vol. 14, no. 2, pp. 3105–3118, 2025.
- [2] A. W. Anggraeni, A. S. Fitriani, and A. Eviyanti, "Penerapan Algoritma Support Vector Machine untuk Memprediksi Tingkat Partisipasi Pemilu terhadap Kualitas Pendidikan," *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 1, pp. 21–27, 2024, doi: 10.29408/edumatic.v8i1.24838.
- [3] Y. Raharja, A. S. Fitriani, and R. Dijaya, "Klasifikasi Tingkat Partisipasi Pemilu Berdasarkan Sektor Industri Menggunakan Algoritma Naïve Bayes," *Jurnal Tekinkom (Teknik Informasi dan Komputer)*, vol. 7, no. 1, pp. 135–143, 2024, doi: 10.37600/tekinkom.v7i1.1204.
- [4] M. J. Siddiq et al., "ANALISIS SENTIMEN MASYARAKAT TERHADAP PILKADA 2024 DI MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA NAÏVE," vol. 13, no. 2, pp. 609–622, 2025.
- [5] D. E. Safitri and A. S. Fitriani, "Implementasi Metode Klasifikasi Dengan Algoritma Support Vector Machine Kernel Gaussian Rbf Untuk Prediksi Partisipasi Pemilu Terhadap Demografi Kota Surabaya," *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 5, no. 1, p. 36, 2022, doi: 10.21927/ijubi.v5i1.2259.
- [6] A. Akram et al., "MINI-BATCH K-MEANS CLUSTERING UNTUK PENGELOMPOKAN," pp. 235–244, 2024.
- [7] Y. Izamahendra and M. Subhan, "DESA PARIAMAN TIMUR BERDASARKAN INDEKS DESA MEMBANGUN 2023," vol. 13, no. 2, pp. 204–211, 2025.
- [8] A. Sentimen, A. Pemilu, and A. N. Bayes, "SWADHARMA (JEIS)," vol. 05, pp. 131–139, 2025.

- [9] A. S. Kirana, R. Roeswidiyah, and A. Pudoli, "ANALISIS SENTIMEN PADA MEDIA SOSIAL TERHADAP LAYANAN SAMSAT DIGITAL NASIONAL," vol. 8, pp. 53–63, 2025.
- [10] D. M. Chulloh, A. S. Fitriani, I. R. Indra Astutik, and A. Eviyanti, "Uji Akurasi K-Means dalam Prediksi Partisipasi Pemilu pada Demografi Wilayah Kabupaten Pasuruan," *Jutisi : Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 13, no. 1, p. 201, 2024, doi: 10.35889/jutisi.v13i1.1753.
- [11] R. A. Kurniawan, F. I. N. G, and M. R. Pribadi, "Sentimen Analisis Presiden Terpilih Menggunakan Algoritma Support Vector Machine," *The Indonesian Journal of Computer Science Research*, vol. 3, no. 14, pp. 68–73, 2024.
- [12] E. F. Laili et al., "KOMPARASI ALGORITMA DECISION TREE DAN SUPPORT VECTOR MACHINE (SVM) DALAM," vol. 8, no. 1, pp. 67–76, 2025.
- [13] D. E. Safitri and A. S. Fitriani, "Implementasi Metode Klasifikasi Dengan Algoritma Support Vector Machine Kernel Gaussian Rbf Untuk Prediksi Partisipasi Pemilu Terhadap Demografi Kota Surabaya," *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 5, no. 1, p. 36, 2022, doi: 10.21927/ijubi.v5i1.2259.
- [14] A. W. Anggraeni, A. S. Fitriani, and A. Eviyanti, "Penerapan Algoritma Support Vector Machine untuk Memprediksi Tingkat Partisipasi Pemilu terhadap Kualitas Pendidikan," *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 1, pp. 21–27, 2024, doi: 10.29408/edumatic.v8i1.24838.
- [15] A. S. Kirana, R. Roeswidiyah, and A. Pudoli, "ANALISIS SENTIMEN PADA MEDIA SOSIAL TERHADAP LAYANAN SAMSAT DIGITAL NASIONAL," vol. 8, pp. 53–63, 2025.

Conflict of Interest Statement:

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.