

OPTIMASI PENINGKATAN AKURASI KLASIFIKASI STUNTING MENGGUNAKAN ALGORITHM SUPPORT VECTOR MACHINE

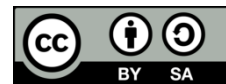
I Gusti Prahmana¹, Kristina Annatasia Br Sitepu², Adek Maulidya³, Toni Prabowo⁴

1, 2, 3) Teknologi Informasi & Informatika, Ilmu Komputer, STMIK Kaputama, Indonesia

4) Teknologi Informasi, Universitas Pembangunan Panca Budi, Indonesia

Article Info	ABSTRACT
Article history: Received: 08 Juli 2026 Revised: 13 Juli 2026 Accepted: 16 Juli 2026	Abstrak Stunting merupakan tantangan kesehatan masyarakat yang berdampak signifikan pada perkembangan fisik dan kognitif anak di masa depan. Penelitian ini bertujuan untuk mengoptimalkan akurasi klasifikasi status gizi balita menggunakan algoritma <i>Support Vector Machine</i> (SVM). Dataset yang digunakan terdiri dari 161 data sekunder dari Dinas Pengendalian Penduduk dan Keluarga Berencana (PPKB) Kota Binjai, mencakup atribut umur, tinggi badan, dan berat badan. Metodologi penelitian meliputi tahap praproses, seleksi fitur, serta optimasi model melalui penyetelan parameter kernel, nilai <i>cost</i> (C), dan <i>gamma</i>. Evaluasi kinerja dilakukan menggunakan <i>confusion matrix</i> dengan metrik akurasi, presisi, <i>recall</i>, dan <i>F1-score</i>. Hasil penelitian menunjukkan bahwa model SVM yang dioptimasi berhasil mencapai tingkat akurasi sebesar 84,85%. Optimasi pada penelitian ini dilakukan untuk meningkatkan kemampuan algoritma <i>Support Vector Machine</i> (SVM) dalam mengklasifikasikan status stunting secara lebih akurat. Oleh karena itu, diperlukan proses optimasi agar model menghasilkan performa yang lebih optimal. Model ini menunjukkan performa yang sangat kuat dalam mengidentifikasi status gizi normal, meskipun masih terdapat kendala dalam mengklasifikasikan kategori kurus akibat ketidakseimbangan data. Secara keseluruhan, algoritma SVM terbukti efektif dalam memetakan risiko stunting secara dini. Temuan ini diharapkan dapat menjadi instrumen pendukung bagi otoritas kesehatan dalam merancang program intervensi yang lebih presisi dan tepat sasaran bagi masyarakat. Kata Kunci: Stunting, Support Vector Machine, Optimasi, Machine Learning, Akurasi. Abstract <i>Stunting is a public health challenge that significantly impacts children's future physical and cognitive development. This study aims to optimize the accuracy of toddler nutritional status classification using the Support Vector Machine (SVM) algorithm. The dataset used consists of 161 secondary data from the Population and Family Planning Control Office (PPKB) of Binjai City, including attributes of age, height, and weight. The research methodology includes preprocessing, feature selection, and model optimization through kernel parameter tuning, cost (C) values, and gamma. Performance evaluation was conducted using a confusion matrix with metrics of accuracy, precision, recall, and F1-score. The results showed that the optimized SVM model successfully achieved an accuracy level of 84.85%. Optimization in this study was carried out to improve the ability of the Support Vector Machine (SVM) algorithm to classify stunting status more accurately. Therefore, an optimization process is needed for the model to produce more optimal performance. This model showed very strong performance in identifying normal nutritional status, although there were still obstacles in classifying the thin category due to data imbalance. Overall, the SVM algorithm proved effective in early mapping the risk of stunting. These findings are expected to be a supporting instrument for health authorities in designing more precise and targeted intervention programs for the community.</i> Keywords: Stunting, Support Vector Machine, Optimization, Machine Learning, Accuracy.

didistribusikan di bawah syarat dan ketentuan dengan Lisensi Internasional Creative Commons Attribution NonCommercial ShareAlike 4.0 ([CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/)).



Corresponding Author: I Gusti Prahmana

E-mail : igustiprahmana4@gmail.com

1. PENDAHULUAN

Sebagai salah satu negara berkembang, Indonesia masih menghadapi tantangan besar terkait masalah gizi kronis, khususnya stunting. Masalah ini dipicu oleh defisit nutrisi berkepanjangan pada fase 1.000 Hari Pertama Kehidupan (HPK), yang secara klinis diidentifikasi dari tinggi badan anak yang berada di bawah standar usianya. Dampak dari stunting tidak hanya terbatas pada hambatan pertumbuhan fisik secara jangka pendek, melainkan juga berisiko menurunkan kapasitas kognitif, menghambat produktivitas kerja, serta meningkatkan kerentanan terhadap penyakit tidak menular saat dewasa [1, 9]. Oleh karena itu, percepatan reduksi angka stunting diintegrasikan ke dalam program prioritas nasional guna mencetak generasi masa depan yang lebih sehat dan berdaya saing.

Di tingkat daerah, Kota Binjai secara aktif mengimplementasikan program penurunan stunting lewat kolaborasi lintas sektor yang melibatkan pemerintah daerah, Dinas Pengendalian Penduduk dan Keluarga Berencana (PPKB), pemangku fasilitas kesehatan, serta elemen masyarakat. Berbagai langkah taktis telah digulirkan, meliputi Program Bangga Kencana, Bapak Asuh Anak Stunting (BAAS), edukasi nutrisi, hingga pemberian makanan tambahan. Meski demikian, penurunan angka kasus di lapangan masih menemui jalan buntu. Tingginya prevalensi stunting yang belum mencapai target nasional dipengaruhi oleh beberapa kendala struktural, seperti keterbatasan sumber daya, rendahnya partisipasi publik, dan lemahnya koordinasi antarinstansi [10].

Hasil pemetaan dari Dinas PPKB Kota Binjai memperlihatkan bahwa pola persebaran keluarga rentan stunting tersebar di berbagai kecamatan dengan karakteristik wilayah yang variatif. Munculnya risiko ini dipengaruhi oleh faktor-faktor multidimensi, mulai dari aspek lingkungan, sanitasi, tingkat pendidikan keluarga, kondisi ekonomi, hingga kesehatan reproduksi [11]. Menurut World Health Organization (WHO), stunting merupakan indikator utama dalam mengukur defisiensi gizi kronis pada anak yang mencerminkan ketidakseimbangan kronis antara asupan dan kebutuhan nutrisi [2]. Walaupun pemerintah telah melakukan intervensi multi-sektor, proses deteksi dini

kelompok berisiko masih terhambat oleh kompleksitas dan besarnya volume data di lapangan, sehingga diperlukan model pengolahan data yang lebih taktis dan akurat.

Pemanfaatan teknologi informasi dan kecerdasan buatan, khususnya *machine learning*, membuka peluang baru dalam mengatasi problem data medis ini melalui standarisasi sistem klasifikasi otomatis [3]. Dalam penanganan stunting, *machine learning* dapat dioptimalkan untuk memetakan variabel prediktif sekaligus mengelompokkan status gizi anak secara cepat. Salah satu algoritma klasifikasi yang dikenal memiliki performa tinggi adalah *Support Vector Machine* (SVM). Algoritma ini bekerja dengan menentukan *hyperplane* paling optimal untuk memisahkan data antarkelas dengan margin yang maksimal [4]. Keunggulan utama SVM terletak pada efektivitasnya dalam memproses data berdimensi tinggi serta kemampuannya meminimalkan risiko *overfitting*, menjadikannya sangat relevan untuk analisis data klinis.

Akan tetapi, performa klasifikasi SVM sangat bergantung pada penentuan parameter dan sifat data yang dianalisis. Konfigurasi parameter yang kurang tepat berisiko menurunkan akurasi model secara drastis, sehingga diperlukan proses tuning parameter untuk mendapatkan kombinasi terbaik [5]. Pendekatan optimasi seperti *Grid Search*, *Random Search*, atau metode optimasi lainnya dapat diterapkan guna mencari nilai parameter *kernel*, *cost* (C), dan *gamma* yang paling adaptif dengan karakteristik data stunting.

Berdasarkan urgensi tersebut, penelitian ini difokuskan untuk meningkatkan akurasi model klasifikasi stunting melalui optimasi algoritma *Support Vector Machine*. Melalui komparasi dan pencarian parameter yang presisi, model yang dibangun diharapkan mampu memberikan tingkat akurasi yang lebih tinggi dalam mendeteksi status stunting anak. Luaran dari penelitian ini diproyeksikan menjadi solusi berbasis teknologi yang dapat mendukung pengambilan keputusan klinis secara lebih cepat, tepat sasaran, dan efisien bagi pemangku kebijakan.

Stunting masih menjadi salah satu permasalahan kesehatan masyarakat yang memerlukan metode klasifikasi dengan tingkat akurasi tinggi untuk mendukung deteksi dini. Berbagai penelitian telah menerapkan algoritma *Machine Learning* seperti Naïve Bayes, Decision Tree, Random Forest, Artificial Neural Network (ANN), dan Support Vector Machine (SVM) dalam klasifikasi status stunting. Meskipun SVM dikenal memiliki

kemampuan generalisasi yang baik pada data berukuran kecil hingga menengah serta efektif dalam menangani data berdimensi tinggi, sebagian besar penelitian sebelumnya masih menggunakan konfigurasi parameter *default*. Kondisi tersebut menyebabkan performa model belum mencapai tingkat optimal karena sangat dipengaruhi oleh pemilihan parameter seperti kernel, cost (C), dan gamma.

Selain itu, penelitian terdahulu umumnya hanya berfokus pada pelaporan nilai akurasi tanpa melakukan proses optimasi parameter secara sistematis. Padahal, performa SVM sangat sensitif terhadap kombinasi parameter yang digunakan. Di sisi lain, evaluasi model pada penelitian sebelumnya juga masih terbatas pada metrik akurasi, sementara metrik lain seperti precision, recall, F1-score, dan Area Under Curve (AUC) sering kali belum dianalisis secara komprehensif.

Penelitian lain juga menunjukkan bahwa data stunting umumnya memiliki karakteristik berupa ketidakseimbangan kelas (*class imbalance*), keberadaan atribut yang saling berkorelasi, serta distribusi data yang tidak linier.

2. METODE PENELITIAN

Tahapan dalam penelitian ini dilakukan secara sistematis mengacu pada alur kerja yang direpresentasikan pada Gambar 1. Kerangka kerja tersebut disusun secara terstruktur untuk membangun model klasifikasi stunting menggunakan algoritma *Support Vector Machine* (SVM) yang terintegrasi dengan metode optimasi parameter. Setiap fase di dalam alur penelitian ini memiliki fungsionalitas dan target spesifik, mulai dari pengolahan awal data hingga tahap pengujian, guna menjamin model klasifikasi yang dihasilkan memiliki tingkat akurasi dan performa yang optimal.



Gambar 1. Alur Sistematis Pemodelan Klasifikasi Stunting Menggunakan Algoritma *Support Vector Machine* (SVM) yang Dioptimasi

2.1 Pengumpulan Data

Data dalam penelitian ini diperoleh melalui kolaborasi dengan Dinas Pengendalian Penduduk dan Keluarga Berencana (PPKB) Kota Binjai, selaku instansi berwenang dalam program akselerasi reduksi stunting di tingkat daerah. Jenis data yang digunakan adalah data sekunder, yang bersumber dari rekaman observasi lapangan oleh kader dan tenaga kesehatan saat melakukan pemantauan berkala terhadap pertumbuhan balita.

Dataset yang berhasil dihimpun mencakup 161 rekam data balita dengan lima variabel utama, yaitu umur (tahun), umur (bulan), tinggi badan (cm), berat badan (kg), serta status gizi. Variabel status bertindak sebagai label atau target keluaran (*class target*), yang merepresentasikan klasifikasi klinis antropometri anak ke dalam dua kategori, yakni "Normal" dan "Stunting". Struktur dan ringkasan atribut dari dataset yang digunakan dalam eksperimen ini disajikan pada Tabel 1.

Sebelum diintegrasikan ke dalam arsitektur algoritma klasifikasi, dataset mentah ini melewati tahapan *data checking* secara menyeluruh. Proses ini bertujuan untuk mengidentifikasi sekaligus mengeliminasi adanya nilai yang hilang (*missing value*), data ganda (*duplicate data*), maupun inkonsistensi format penulisan. Implementasi praproses data (*preprocessing*) ini krusial untuk mentransformasikan seluruh fitur numerik ke dalam format yang kompatibel, sehingga kualitas data meningkat dan mampu mengoptimalkan akurasi model klasifikasi yang dibangun.

Variabel usia secara spesifik dipisahkan ke dalam satuan tahun dan bulan guna memperoleh representasi rentang umur balita yang lebih presisi. Di sisi lain, parameter tinggi badan dan berat badan diposisikan sebagai indikator utama dalam mengukur deviasi pertumbuhan fisik anak. Integrasi dari keempat fitur masukan tersebut menjadi fondasi utama dalam fase pelatihan (*training*) model menggunakan algoritma *Support Vector Machine* (SVM).

Dataset dari Dinas PPKB Kota Binjai ini memiliki tingkat validitas dan representasi yang tinggi karena dikumpulkan langsung oleh tenaga kompeten di lapangan. Melalui pemanfaatan data ini, model prediksi yang dihasilkan diharapkan dapat memberikan kontribusi nyata bagi pemerintah daerah sebagai instrumen deteksi dini yang cepat, akurat, dan efektif dalam memetakan balita berisiko stunting.

Tabel 1. Sampel Dataset Pertumbuhan Balita dari Dinas PPKB Kota Binjai

No	Umur Tahun	Umur Bulan	Tinggi	Berat Badan	Status
1	4	6	104,8	16,5	Normal
2	4	2	104,5	15,9	Normal
3	2	3	92,4	13,6	Normal
4	1	8	85	12,7	Gemuk
5	4	4	104,5	16,9	Normal
6	5	0	107,3	17,2	Normal
7	3	3	102,3	15,2	Normal
8	3	4	101,5	15,4	Normal
9	3	4	102,2	15,6	Normal
10	4	8	106,3	16,9	Normal
11	0	3	58,5	5,7	Normal
12	2	4	92,5	13,9	Normal
13	3	6	103,5	16	Normal
14	1	8	87,5	12,9	Normal
15	1	7	86,8	12,8	Normal
16	1	9	85,9	12,7	Gemuk
17	4	6	105,8	16,5	Normal
18	3	7	101,3	16	Normal
19	0	6	67	7,6	Normal
20	3	8	101	15,8	Normal
21	2	0	93	13,8	Normal
22	1	1	81	11,5	Gemuk
23	4	4	106,3	16,7	Normal
24	4	4	105	16,6	Normal
25	3	2	98,3	15,5	Normal
26	4	4	104,6	17,2	Normal
27	1	11	87,8	13,1	Normal
28	1	9	89	13,2	Normal
29	0	4	62	6,4	Normal
30	0	9	72,8	10,8	Gemuk
31	0	10	73,5	11	Gemuk
32	1	5	83,5	12,3	Gemuk
33	1	10	88	12,5	Normal
34	1	3	81,8	11,7	Gemuk
35	1	10	89,7	13,4	Normal
36	1	10	90	13,2	Normal
37	3	9	103,5	15,9	Normal
38	3	9	103,5	15,9	Normal
39	1	3	82	11,7	Gemuk
40	1	3	82	11,7	Gemuk
41	2	10	95	14,5	Normal

No	Umur Tahun	Umur Bulan	Tinggi	Berat Badan	Status
42	2	4	93,4	13,8	Normal
43	2	2	92,5	13,4	Normal
44	1	4	83	12,3	Gemuk
45	1	4	83	12,3	Gemuk
46	2	7	94,5	14	Normal
47	1	4	82,6	11,9	Gemuk
48	3	4	99	15,8	Normal
49	0	11	78,5	11,2	Gemuk
50	0	9	74	10,5	Gemuk
51	0	10	76	10,3	Gemuk
52	3	6	104	15,8	Normal
53	1	4	83	11,5	Normal
54	3	7	101,9	15,5	Normal
55	1	5	83,4	11,4	Normal
56	3	6	96	14,9	Normal
57	1	3	82,8	11,8	Gemuk
58	2	2	89	13,2	Normal
59	2	2	89	13,2	Normal
60	2	10	94,5	14,7	Normal
61	4	5	105,4	16,1	Normal
62	3	11	103	15,9	Normal
63	4	10	107,2	17,2	Normal
64	0	8	71	9,9	Gemuk
65	0	11	78	11,4	Gemuk
66	1	7	86	13,2	Gemuk
67	1	1	80	10,9	Gemuk
68	3	1	97,7	15,2	Normal
69	1	8	86,7	13	Gemuk
70	4	3	105,5	16,2	Normal
71	4	4	104,2	16,3	Normal
72	1	7	84,6	10,1	Normal
73	2	7	93	14	Normal
74	2	5	93	13,6	Normal
75	1	0	75,4	11,6	Gemuk
76	1	6	85,8	12,8	Gemuk
77	3	10	102,5	14,9	Normal
78	3	2	98,5	14,8	Normal
79	4	10	107	17	Normal
80	1	7	91,4	12,5	Normal
81	1	2	80,5	11,7	Gemuk
82	1	7	83,3	11,9	Gemuk
83	1	0	76,4	11,4	Gemuk
84	1	6	87,2	12,8	Normal

No	Umur Tahun	Umur Bulan	Tinggi	Berat Badan	Status
85	2	5	90,2	13,3	Normal
86	4	3	103,7	15,6	Normal
87	1	8	88,3	13,2	Normal
88	4	8	105,5	16,5	Normal
89	4	6	104,2	15,4	Normal
90	1	7	85	12,4	Gemuk
91	3	6	99,2	15	Normal
92	2	8	94	13,9	Normal
93	1	10	87,5	13,4	Gemuk
94	4	6	104,3	16,6	Normal
95	4	6	106,5	17,8	Normal
96	0	8	71	9,8	Gemuk
97	1	7	86,5	13,2	Gemuk
98	1	2	82,8	11,8	Gemuk
99	3	4	99,4	15,3	Normal
100	0	9	74,5	10,9	Gemuk
101	4	11	107	18	Normal
102	2	2	89	13,3	Normal
103	3	7	101,7	15,6	Normal
104	2	0	90,6	13,7	Normal
105	0	9	69	10,3	Gemuk
106	0	11	75	11,2	Gemuk
107	1	2	82,6	11,5	Normal
108	1	5	79,5	9,9	Normal
109	3	11	101,7	14,9	Normal
110	4	0	100	15	Normal
111	4	9	104,5	16,2	Normal
112	1	9	88,4	12,5	Normal
113	1	9	86	12,5	Normal
114	1	4	82	12	Gemuk
115	1	6	84,5	12,3	Gemuk
116	2	0	92	13,4	Normal
117	0	9	73	10,7	Gemuk
118	3	2	99,5	15,2	Normal
119	0	6	63	7	Gemuk
120	3	1	98	14,8	Normal
121	0	10	73	10,9	Gemuk
122	3	3	98,4	14,9	Normal
123	4	5	105,8	16	Normal
124	1	8	88	13,2	Gemuk
125	1	4	82,5	12,1	Gemuk
126	1	4	82,5	12,1	Gemuk
127	2	9	94,5	13,9	Normal

No	Umur Tahun	Umur Bulan	Tinggi	Berat Badan	Status
128	1	10	86,5	13,2	Gemuk
129	0	11	75	10,6	Gemuk
130	0	8	68	9,1	Gemuk
131	0	5	64,5	7	Normal
132	4	9	105,3	16,3	Normal
133	4	8	105,3	16,5	Normal
134	2	3	84,5	11,9	Normal
135	2	2	83,9	9	Kurus
136	2	4	92,7	13	Normal
137	1	7	84	11,9	Normal
138	0	5	62	6,4	Normal
139	4	3	106	16,7	Normal
140	5	0	106,7	17,8	Normal
141	0	10	74	11,1	Gemuk
142	2	7	94	14,5	Normal
143	3	3	99,5	14,7	Normal
144	3	3	100	15,1	Normal
145	2	1	90	13,8	Gemuk
146	4	0	106	15,4	Kurus
147	4	0	106	15,3	Kurus
148	0	6	66	8,2	Gemuk
149	0	7	68	8,6	Gemuk
150	4	6	106,8	16,6	Normal
151	1	11	87,5	13	Normal
152	1	3	83	12,2	Gemuk
153	1	10	86	13,1	Gemuk
154	2	10	95,5	14,8	Normal
155	1	10	87,5	13,1	Gemuk
156	3	3	100	14,9	Normal
157	2	5	93,2	13,9	Normal
158	1	0	77	10,9	Gemuk
159	1	5	85,8	13,9	Gemuk
160	5	0	106,5	17,4	Normal
161	4	5	104,5	15,9	Normal

Implementasi sistem dalam eksperimen ini dibangun menggunakan Google Colaboratory (Google Colab) sebagai lingkungan pengembangan berbasis *cloud computing*. Pemilihan Google Colab didasarkan pada efisiensi akses komputasi tingkat tinggi yang ditawarkannya tanpa mengharuskan instalasi perangkat lunak lokal pada perangkat keras pengguna. Di samping itu, platform ini telah terintegrasi secara bawaan dengan ekosistem pustaka (*library*) Python untuk kebutuhan sains data dan *machine*

learning, termasuk Pandas untuk manipulasi data, NumPy untuk operasi matriks, Scikit-learn untuk pemodelan prediktif, serta Matplotlib untuk visualisasi data.

Berdasarkan hasil eksperimen, Google Colab terbukti andal dalam memfasilitasi seluruh rangkaian proses penelitian, yang mencakup manipulasi awal dataset (*preprocessing*), konstruksi arsitektur *Support Vector Machine* (SVM), hingga kalkulasi metrik evaluasi klasifikasi. Penggunaan platform berbasis awan ini tidak hanya mengoptimalkan efisiensi waktu komputasi, tetapi juga mendukung prinsip *reproducible research*, sehingga memudahkan peneliti lain untuk mereplikasi atau memvalidasi kembali tahapan pemodelan yang telah dilakukan.

2.3 Metode Klasifikasi SVM

Support Vector Machine (SVM) merupakan salah satu algoritma *supervised machine learning* yang diimplementasikan secara luas guna menyelesaikan problem klasifikasi maupun regresi. Dalam melakukan klasifikasi, prinsip kerja SVM bertumpu pada pencarian *hyperplane* optimal yang berfungsi sebagai pemisah antarkelas data dengan memaksimalkan nilai margin terhadap objek terdekat dari tiap-tiap kelas. Karakteristik ini membekali SVM dengan kemampuan generalisasi yang tinggi sekaligus performa akurasi yang unggul, khususnya saat memproses dataset berdimensi tinggi meskipun ketersediaan sampel relatif terbatas. Berbagai literatur mutakhir mengonfirmasi bahwa SVM tetap menjadi algoritma yang kompetitif di pelbagai domain terapan, seperti sektor kesehatan, pengolahan citra, bioinformatika, keamanan siber, hingga *Internet of Things* (IoT) [17].

Prinsip dasar dari SVM adalah menetapkan batas pemisah (*hyperplane*) terbaik yang mampu memberikan jarak (margin) paling renggang antar-dua kelas data. Pembentukan *hyperplane* ini ditentukan secara langsung oleh beberapa sampel data latih yang posisinya paling dekat dengan garis pembatas. Sampel-sampel spesifik inilah yang disebut sebagai *support vectors*. Optimalnya margin yang terbentuk berbanding lurus dengan kemampuan generalisasi model terhadap data baru, yang secara simultan dapat menekan risiko terjadinya gejala *overfitting* [14].

Secara matematis, persamaan *hyperplane* pemisah pada ruang linear dirumuskan sebagai berikut:

$$w^T x + b = 0$$

Dalam hal ini, w merepresentasikan vektor bobot (*weight vector*), x melambangkan vektor fitur masukan, dan b menyatakan nilai bias (*intercept*). Proses penentuan kelas dari data uji baru dieksekusi melalui fungsi keputusan (*decision function*) berikut:

$$f(x) = \text{sign}(w^T x + b)$$

Apabila kalkulasi fungsi menghasilkan nilai $f(x) > 0$, objek tersebut akan dimasukkan ke dalam kategori kelas positif (+1). Sebaliknya, jika $f(x) < 0$, objek dikelompokkan ke dalam kelas negatif (-1). Kehadiran fungsi keputusan ini memungkinkan proses klasifikasi berjalan secara efisien untuk mengenali orientasi data baru [17].

Untuk mengonstruksi *hyperplane* dengan margin yang maksimal, SVM menyelesaikan formula fungsi tujuan optimasi konveks sebagai berikut:

$$\min \frac{1}{2} \| w \|^2$$

Persamaan di atas diselesaikan dengan memperhatikan batasan (*constraints*) berupa:

$$y_i(w^T x_i + b) \geq 1, i = 1, 2, \dots, n$$

Keterangan variabel penjelas:

- a. x_i = data ke- i
- b. y_i = label kelas (+1 atau -1)
- c. n = jumlah data pelatihan

Fungsi objektif tersebut bekerja dengan meminimalkan nilai norma vektor bobot demi tercapainya margin antarkelas yang maksimal. Pendekatan optimasi berbasis batas linear inilah yang menjadi distingsi utama SVM dibandingkan dengan algoritma klasifikasi konvensional lain, seperti *Decision Tree* maupun *K-Nearest Neighbor* [15].

Pada skenario data dunia nyata, sebaran data umumnya bersifat non-linear (*non-linearly separable*). Guna mengatasi batasan tersebut, SVM mengadopsi mekanisme *kernel trick*, yaitu sebuah metode pemetaan ruang fitur awal ke ruang spasial berdimensi lebih tinggi sehingga data dapat dipisahkan secara linear melalui *hyperplane*. Teknik ini mengizinkan sistem untuk memproses pola data yang rumit tanpa perlu mengeksplisitkan kalkulasi transformasi fitur berulang kali. Oleh karena itu, pemilihan fungsi kernel menjadi parameter penentu yang krusial bagi akurasi klasifikasi SVM [17]. Beberapa tipe fungsi kernel yang sering diimplementasikan dalam pemodelan meliputi:

1. Kernel Linear

$$K(x_i, x_j) = x_i^T x_j$$

Fungsi ini diterapkan apabila sebaran data secara natural dapat dipisahkan secara linear tanpa perlu transformasi dimensi yang tinggi.

2. Kernel Polynomial

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d$$

Kernel polinomial digunakan untuk membentuk batas keputusan (*decision boundary*) yang dinamis sesuai dengan derajat/orde (d) polinomial yang dikonfigurasi.

3. Kernel *Radial Basis Function* (RBF)

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

RBF merupakan fungsi kernel yang paling populer karena memiliki fleksibilitas tinggi dalam memetakan relasi non-linear pada beragam karakteristik data.

4. Kernel Sigmoid

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$$

Karakteristik kernel sigmoid menyerupai model komputasi fungsi aktivasi yang jamak ditemukan pada arsitektur Jaringan Saraf Tiruan [14].

Di samping penentuan jenis kernel, tingkat akurasi SVM dipengaruhi oleh regulasi parameter esensial, utamanya parameter C dan γ (gamma). Parameter C memegang kendali atas titik keseimbangan (*trade-off*) antara pengejaran margin maksimum dan toleransi kesalahan klasifikasi pada data latih. Nilai C yang tinggi akan mempersempit margin demi meminimalkan galat klasifikasi, sedangkan nilai C yang rendah akan melonggarkan batas margin dengan konsekuensi toleransi kesalahan yang lebih longgar [12].

Secara umum, keunggulan performa SVM terletak pada efektivitasnya memproses dimensi data yang padat, ketahanan terhadap *overfitting*, kemampuan generalisasi yang andal, serta stabilitas akurasi yang konsisten meski dilatih menggunakan jumlah sampel yang terbatas [12].

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian Model

Eksperimen pemodelan dalam penelitian ini diwujudkan menggunakan bahasa pemrograman Python dengan memanfaatkan platform Google Colaboratory (Google Colab) sebagai lingkungan pengembangan (*development environment*) berbasis komputasi awan. Pemilihan platform ini didasarkan pada efisiensinya dalam menyediakan ekosistem kerja siap pakai yang telah terintegrasi dengan berbagai pustaka (*library*) esensial untuk *machine learning*, seperti Pandas, NumPy, dan Scikit-learn, tanpa mengharuskan konfigurasi perangkat lunak secara lokal.

Secara teknis, tahapan implementasi dijalankan melalui alur kerja terstruktur yang diawali dengan pemanggilan (*importing*) pustaka-pustaka yang diperlukan, dilanjutkan dengan pemuatan (*loading*) dataset balita ke dalam memori kerja. Langkah berikutnya adalah melakukan pemisahan antara variabel fitur (*features*) sebagai masukan prediktif dan variabel label (*target*) sebagai output klasifikasi. Dataset kemudian dipartisi menjadi data latih (*training data*) dan data uji (*testing data*) dengan proporsi yang telah ditentukan.

Guna menyamakan skala nilai antar-atribut agar proses konvergensi algoritma berjalan lebih cepat, dilakukan standarisasi data menggunakan metode StandardScaler. Setelah data berada pada skala yang seragam, model *Support Vector Machine* (SVM) dikonstruksi dengan menerapkan fungsi kernel *Radial Basis Function* (RBF). Fase akhir dari implementasi ini adalah melakukan pengujian prediksi pada data uji serta menghitung metrik evaluasi untuk mengukur performa model. Seluruh rangkaian instruksi komputasi ini dieksekusi secara sistematis menggunakan pustaka Scikit-learn di dalam *environment* Google Colab.

3.2 Analisis Hasil Pengujian Model

Berdasarkan eksperimen yang telah dilakukan, model *Support Vector Machine* (SVM) dengan kernel RBF menghasilkan tingkat akurasi sebesar 84,85%. Capaian ini mengindikasikan bahwa arsitektur model memiliki kemampuan yang andal dalam mengklasifikasikan status gizi balita, di mana sekitar 85 dari 100 sampel data uji berhasil diprediksi secara tepat sesuai dengan kondisi riil di lapangan. Hasil ini sekaligus membuktikan validitas fungsi kernel dalam memetakan korelasi non-linear antara

variabel masukan (umur tahun, umur bulan, tinggi badan, dan berat badan) terhadap variabel target status gizi.

Classification Report				
	precision	recall	f1-score	support
Gemuk	1.00	0.64	0.78	11
Kurus	0.00	0.00	0.00	1
Normal	0.81	1.00	0.89	21
accuracy			0.85	33
macro avg	0.60	0.55	0.56	33
weighted avg	0.85	0.85	0.83	33

Gambar 2. Hasil Classification Report Model SVM

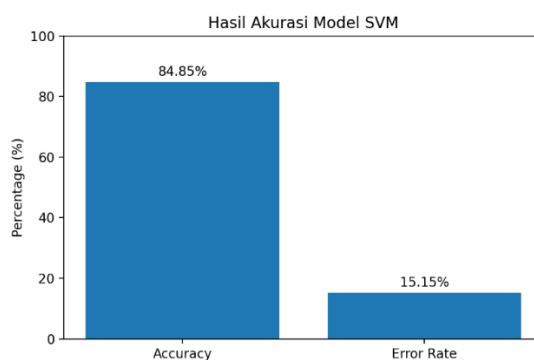
Rincian performa klasifikasi untuk setiap label target disajikan secara spesifik melalui laporan klasifikasi (*classification report*) pada Gambar 2. Metrik evaluasi mencakup pengukuran terhadap *precision*, *recall*, *F1-score*, dan *support* untuk masing-masing kategori status gizi, yaitu Gemuk, Kurus, dan Normal.

Pada kategori status "Gemuk" dengan bobot *support* sebanyak 11 sampel, model mencatatkan nilai *precision* sempurna sebesar 1,00, nilai *recall* sebesar 0,64, dan *F1-score* sebesar 0,78. Perolehan skor *precision* yang menyentuh angka 1,00 mengonfirmasi bahwa seluruh sampel yang diidentifikasi oleh sistem sebagai kelas Gemuk terbukti valid dan akurat (nol persentase *false positive*). Kendati demikian, skor *recall* yang berada di angka 0,64 mengindikasikan keterbatasan model yang hanya mampu mendeteksi 64% dari total data aktual kelas Gemuk, sementara sisanya mengalami salah klasifikasi (*misclassification*) ke kelas lain. Nilai *F1-score* sebesar 0,78 menunjukkan titik keseimbangan performa yang sudah cukup akurat pada kelompok ini, walaupun aspek sensitivitas deteksinya masih memerlukan optimasi lebih lanjut.

Kondisi kontras terjadi pada klasifikasi kelas "Kurus", di mana seluruh metrik pengujian (*precision*, *recall*, dan *F1-score*) bernilai 0,00 dengan jumlah *support* yang hanya bermuatan 1 sampel data. Hasil minor ini dipicu oleh adanya fenomena ketidakseimbangan kelas (*class imbalance*) yang ekstrem pada dataset. Keterbatasan representasi data pada kelas Kurus mengakibatkan algoritma tidak mendapatkan asupan informasi yang memadai selama fase pelatihan untuk mempelajari karakteristik geometris kelas tersebut, sehingga model gagal melakukan generalisasi saat fase pengujian.

Sebaliknya, pada kelas "Normal" dengan volume *support* terbesar sebanyak 21 sampel, model menunjukkan performa superior dengan capaian *precision* sebesar 0,81,

recall sebesar 1,00, dan *F1-score* sebesar 0,89. Nilai *recall* yang mencapai 1,00 menegaskan bahwa sistem berhasil mengenali seluruh data balita berstatus Normal secara mutlak tanpa adanya galat *false negative*. Sementara itu, nilai *precision* sebesar 0,81 menunjukkan bahwa 81% dari total prediksi kelas Normal bersifat valid, sedangkan 19% sisanya merupakan bias dari kelas lain yang keliru diidentifikasi sebagai kelas Normal. Nilai *F1-score* sebesar 0,89 mengukuhkan bahwa model SVM sangat sensitif dan tangguh dalam mengidentifikasi kelompok balita dengan status gizi Normal.



Gambar 3. Grafik Distribusi Accuracy dan Error Rate Model SVM

Secara agregat, performa keseluruhan sistem diwakili oleh nilai akurasi total sebesar 84,85%, yang bermakna mayoritas sampel pada data uji berhasil diklasifikasikan dengan tepat oleh algoritma SVM. Visualisasi komparasi antara tingkat keberhasilan prediksi (*accuracy*) dan tingkat kegagalan prediksi (*error rate* sebesar 15,15%) dipaparkan pada Gambar 3.

Evaluasi makro (*macro average*) yang menghasilkan nilai *precision* 0,60, *recall* 0,55, dan *F1-score* 0,56 mencerminkan rata-rata performa model secara murni di setiap kelas tanpa menguji pengaruh bobot volume data. Rendahnya skor makro ini disebabkan oleh kegagalan prediksi pada kelas Kurus yang kekurangan sampel. Di sisi lain, perhitungan rata-rata tertimbang (*weighted average*) memberikan hasil yang jauh lebih representatif dengan skor *precision* 0,85, *recall* 0,85, dan *F1-score* 0,83. Hal ini terjadi karena metrik *weighted average* turut memperhitungkan proporsi jumlah data, di mana kelas Normal dan Gemuk mendominasi dataset dan berhasil dieksekusi dengan sangat baik oleh model SVM. Berdasarkan analisis komprehensif ini, dapat disimpulkan bahwa penerapan algoritma SVM memiliki potensi besar untuk dikembangkan sebagai sistem pendukung keputusan deteksi dini status gizi balita.

4. SIMPULAN

Berdasarkan rangkaian eksperimen dan analisis yang telah dipaparkan, dapat disimpulkan bahwa implementasi algoritma *Support Vector Machine* (SVM) memiliki efektivitas yang tinggi dalam mengklasifikasikan status gizi balita sebagai instrumen deteksi dini stunting. Model prediktif yang dikonstruksi menggunakan data sekunder dari Dinas Pengendalian Penduduk dan Keluarga Berencana (PPKB) Kota Binjai ini berhasil menembus tingkat akurasi sebesar 84,85%. Capaian tersebut mengonfirmasi bahwa arsitektur sistem mampu memprediksi sekitar 85 dari 100 data balita secara presisi dan selaras dengan kondisi riil di lapangan.

Hasil evaluasi detail via *classification report* menunjukkan representasi performa yang bervariasi antarkelas. Model menunjukkan superioritas mutlak dalam mengidentifikasi kategori status gizi "Normal" dengan perolehan skor *recall* sempurna senilai 1,00 dan *F1-score* sebesar 0,89. Kendati demikian, eksperimen ini menyingkap keterbatasan sistem dalam mengenali kategori kelas "Kurus". Kendala ini dipicu secara linear oleh anomali ketidakseimbangan data (*class imbalance*) yang ekstrem pada dataset, di mana minimnya kuantitas sampel membuat algoritma kekurangan basis referensi untuk merekam karakteristik geometris kelas tersebut secara optimal.

Secara implikasi praktis, model SVM yang dikembangkan dalam penelitian ini sangat prospektif untuk diadopsi sebagai sistem pendukung keputusan digital oleh otoritas kesehatan di Kota Binjai guna memetakan risiko stunting secara taktis, cepat, dan akurat. Guna menyempurnakan batasan penelitian ini ke depan, pengembangan studi selanjutnya disarankan untuk mengintegrasikan teknik manipulasi data minoritas seperti *Synthetic Minority Over-sampling Technique* (SMOTE) atau metode *undersampling*, serta memperluas cakupan volume dataset agar performa klasifikasi di seluruh kategori status gizi dapat terdistribusi secara lebih merata dan andal.

REFERENCES

- [1] UNICEF, The State of the World's Children 2023: For Every Child, Nutrition. New York, NY, USA: UNICEF, 2023. [Online]. Available: <https://kc.umn.ac.id/id/eprint/40359/>, [Accessed: Apr. 2023].
- [2] World Health Organization, WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age. Geneva, Switzerland: WHO, 2023.
- [3] Tom M. Mitchell, Machine Learning. New York, NY, USA: McGraw-Hill, 1997.
- [4] Corinna Cortes and Vladimir Vapnik, "Support-Vector Networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995. vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.

-
- [5] Kevin P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
 - [6] Kementerian Kesehatan Republik Indonesia, *Buku Saku Hasil Survei Status Gizi Indonesia (SSGI) Tahun 2024*. Jakarta, Indonesia: Kemenkes RI, 2024. Available: <https://www.badankebijakan.kemkes.go.id/survei-status-gizi-indonesia-ssgi-2024>, [Accessed: June. 2024]
 - [7] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009. doi: 10.1007/978-0-387-84858-7.
 - [8] Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
 - [9] Badan Kependudukan dan Keluarga Berencana Nasional, *Panduan Percepatan Penurunan Stunting melalui Program Bangga Kencana*. Jakarta: BKKBN, 2024.
 - [10] Putri, D. A., dan Wiro, "Implementasi Program Bapak Asuh Anak Stunting dalam Penanggulangan Stunting di Kota Binjai," *SAJJANA: Public Administration Review*, vol. 3, no. 2, 2025. vol. 3, no. 2, 2025, doi: 10.32734/sajjana.v3i02.23501.
 - [11] Suraya, R., dkk., "Analisis Spasial Keluarga Risiko Stunting Berdasarkan Jamban Tidak Layak dan PUS 4T di Dinas PPKB Kota Binjai Tahun 2024," *Jurnal Kesehatan Jompa*, vol. 4, no. 4, 2025, 2025, doi: 10.57218/jkj.Vol4.Iss4.2149.
 - [12] Al-Dulaimi, H., & Ku-Mahamud, K. (2025). *A Review on Support Vector Machine Problems and Solutions*. *Data Science Insights*, 4(1). <https://doi.org/10.63017/jdsi.v4i1.217>
 - [13] Kumari, A., Akhtar, M., Shah, R., & Tanveer, M. (2025). *Support Matrix Machine: A Review*. *Neural Networks*, 181, 106767. <https://doi.org/10.1016/j.neunet.2024.106767>
 - [14] Nagar, K., & Chawla, M. P. S. (2023). *A Survey on Various Approaches for Support Vector Machine Based Engineering Applications*. *International Journal of Emerging Science and Engineering*. <https://doi.org/10.35940/IJESE.K2555.1011112>
 - [15] Kim, J., et al. (2023). *A New Support Vector Machine for Categorical Features*. *Expert Systems with Applications*, 229, 120449. <https://doi.org/10.1016/j.eswa.2023.120449>
 - [16] Suharmin, W. N. A., Hasan, I. K., & Yahya, N. I. (2025). *Comparison of Forward Selection and Backward Elimination Feature Selection Methods in Support Vector Machine Algorithm*. *The Indonesian Journal of Computer Science*. <https://doi.org/10.33022/ijcs.v14i2.4755>
 - [17] De Simone, V., et al. (2024). *An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review*. *Information*, 15(4), 235. <https://doi.org/10.3390/info15040235>
 - [18] Wijaya, R. S., Qur'ania, A., & Anggraeni, I. (2024). *Klasifikasi Penyakit Cacar Monyet Menggunakan Support Vector Machine (SVM)*. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*. <https://doi.org/10.57152/malcom.v4i4.1417>
 - [19] Benítez-Peña, S., Blanquero, R., Carrizosa, E., & Ramírez-Cobo, P. (2023). *Cost-Sensitive Probabilistic Predictions for Support Vector Machines*. <https://doi.org/10.48550/arXiv.2310.05997>
 - [20] Rezvani, S., Pourpanah, F., Lim, C. P., & Wu, Q. M. J. (2024). *Methods for Class-Imbalanced Learning with Support Vector Machines: A Review and an Empirical Evaluation*. <https://doi.org/10.48550/arXiv.2406.03398>
 - [21] IC2IE. (2023). *Optimization of SVM Classification Accuracy with Bayesian Optimization Utilizing Data Augmentation*. <https://doi.org/10.1109/IC2IE60547.2023.10331580>