

KLASIFIKASI STATUS GIZI BALITA MENGGUNAKAN *SUPPORT VECTOR MACHINE* (STUDI KASUS : PUSKESMAS OEBOBO KOTA KUPANG)

Maria Putriana Ela, Meliana O. Meo, Skolastika Siba Igon

Teknik Informatika, STIKOM Uyelindo Kupang
Jalan Perintis Kemerdekaan 1, Kayu Putih, Kota Kupang, Indonesia
putriela0804@gmail.com

ABSTRAK

Anak balita usia 0 hingga 59 bulan merupakan kelompok usia yang rentan terhadap masalah gizi sehingga memerlukan pemantauan status gizi yang akurat dan berkelanjutan. Di Puskesmas Oebobo Kota Kupang, proses penentuan status gizi masih dilakukan secara manual, sehingga berpotensi menimbulkan kesalahan pencatatan dan perhitungan. Penelitian ini bertujuan untuk mengklasifikasikan status gizi balita menggunakan metode *Support Vector Machine* (SVM). Metode SVM digunakan karena kemampuannya dalam mencari hyperplane optimal sebagai batas pemisah antar kelas dengan margin maksimum, serta didukung kernel *Radial Basis Function* (RBF) untuk menangani data non-linar. Pengujian dilakukan dengan variasi C dan gamma untuk memperoleh performa model terbaik. Hasil penelitian menunjukkan bahwa kombinasi parameter C 50 dan gamma scale menghasilkan kinerja yang baik dengan akurasi sebesar 80,21%, presisi 79,05%, recall 80,21% dan f1-score 79,18%. Hasil tersebut menunjukkan bahwa model SVM mampu mengklasifikasikan status gizi balita dengan baik dan dapat digunakan sebagai pendukung pengambilan keputusan dalam pemantauan gizi.

Kata kunci : Gizi Balita, Klasifikasi, SVM

1. PENDAHULUAN

Anak balita merupakan kelompok usia 0 hingga 59 bulan yang berada pada fase pertumbuhan dan perkembangan yang sangat pesat[1].

Pada priode usia balita ini sering disebut sebagai masa emas, karena kecukupan asupan gizi yang berkualitas sangat berperan dalam menentukan kualitas kesehatan dan perkembangan anak di masa mendatang[2].

Oleh karena itu, pemantauan status gizi menjadi hal penting untuk mendeteksi dini potensi gangguan pertumbuhan. Berdasarkan data dari Survei Status Gizi Indonesia (SSGI) tahun 2024 yang dirilis oleh Kementerian Kesehatan, presentase kasus stunting di Kota Kupang sebesar 26,9%, balita dengan berat badan kurang dibandingkan dengan standar normal untuk usianya (*Underweight*) sebesar 29,8%, *Wasting* sebesar 12,0%, *Severse Wasting* sebesar 1,8%, dan *Overweight* 1,8% [3].

Data tersebut menunjukkan bahwa permasalahan status gizi masih menjadi isu yang perlu mendapatkan perhatian khusus.

Pemantauan perkembangan balita umumnya dilakukan secara rutin setiap bulan melalui kegiatan Posyandu[1].

Namun, dalam praktiknya masih sering ditemukan kendala seperti kemungkinan terjadi kesalahan dalam proses pencatatan data, perhitungan hasil pengukuran, dan penentuan kategori status gizi. Proses yang masih dilakukan secara manual berpotensi menurunkan tingkat akurasi dan efisiensi dalam pengambilan keputusan.

Untuk mengatasi permasalahan tersebut, diperlukan suatu sistem klasifikasi berbasis teknologi yang membantu proses penentuan status gizi secara lebih akurat. Salah satu metode yang dapat digunakan

adalah *Support Vector Machine* (SVM) dengan kernel RBF. Metode ini dipilih karena kemampuannya secara optimal memisahkan data ke dalam kelas-kelas tertentu, bahkan ketika data tersebut bersifat kompleks dan non-linar seperti dalam mengklasifikasikan status gizi berdasarkan data antropometri yang tidak bisa di pisahkan secara linear[4].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengklasifikasikan status gizi menggunakan *Support Vector Machine* (SVM) kernel RBF. Penerapan metode ini diharapkan dapat membantu meningkatkan ketepatan dalam penentuan status gizi balita serta memberi hasil klasifikasi yang dapat digunakan sebagai pendukung pengambilan keputusan dalam layanan kesehatan.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Ramon *et al* menerapkan Algoritma *Support Vector Machine* (SVM) dalam mengkasifikasikan status gizi bayi di Posyandu Kecamatan Bangun Purba. Penelitian tersebut menggunakan data sebanyak 1.732, dari data tersebut di bagi 80% untuk data latih dan 20%, model yang dihasilkan mampu mencapai akurasi 87,6% dengan nilai F1-score 86,5%[2].

Nadroh *et al* mengembangkan model klasifikasi status gizi balita menggunakan algoritma *Support Vector Machine*. Hasil awal menunjukkan tingkat akurasi sebesar 72,29%, kemudian dilakukan optimasi parameter menggunakan metode *Grid Search Cross-Validation* dengan kernel RBF yang mampu meningkatkan akurasi menjadi 85,71%. Hasil ini menunjukkan bahwa optimasi parameter SVM efektif dalam meningkatkan kinerja model dalam mengklasifikasikan status gizi balita[1].

Candra *et al* mengembangkan model klasifikasi stunting pada balita menggunakan SVM. Penelitian ini menerapkan tahapan prapemrosesan data untuk meningkatkan kualitas model, hasil yang diperoleh menunjukkan bahwa model SVM mampu mencapai tingkat akurasi sebesar 98,37%[5].

2.2. Klasifikasi

Klasifikasi adalah proses membangun model atau fungsi yang mampu membedakan antar kelas data, sehingga dapat digunakan untuk menentukan kategori suatu objek yang belum diketahui. Tujuan klasifikasi adalah mengenali pola atau hubungan tersembunyi dalam data, kemudian memanfaatkan informasi tersebut untuk memprediksi kategori dari data baru yang belum diketahui[6].

2.3. Status Gizi

Status gizi merupakan ukuran keberhasilan dalam pemenuhan kebutuhan gizi seseorang yang biasanya ditunjukkan melalui indikator berat badan dan tinggi badan. Status ini juga mencerminkan kondisi kesehatan yang dihasilkan dari keseimbangan antara kebutuhan dan asupan nutrisi yang diterima[7].

2.4. Machine Learning

Machine learning (ML) merupakan cabang dari kecerdasan buatan (*Artificial Intelligence/AI*) yang berfokus pada pengembangan sistem yang dapat belajar dan meningkatkan kinerja secara otomatis melalui pengalaman, tanpa perlu diprogram secara eksplisit. Dalam penerapannya, algoritma *machine learning* digunakan untuk menganalisis data, mengenali pola, serta membuat Keputusan atau prediksi berdasarkan data yang tersedia. Proses kerja ML dimulai dengan melatih model menggunakan dataset agar dapat memahami hubungan dan pola dalam data tersebut. Setelah proses pelatihan selesai, model mampu memproses data baru dan menghasilkan output berupa prediksi atau klasifikasi

2.5. Support Vector Machine

Support Vector Machine (SVM) adalah mencari hyperplane terbaik yang berfungsi sebagai batas pemisah antara dua kelas pada ruang input[8]. Data atau pola (*pattern*) dalam SVM dikategorikan ke dalam dua kelas, yaitu +1 dan -1, dengan berbagai kemungkinan garis pemisah (*discrimination boundaries*) yang dapat terbentuk. Margin merupakan jarak antara hyperplane dengan data terdekat dari masing-masing kelas, sedangkan data yang berada paling dekat dengan hyperplane tersebut disebut support vector, yang menjadi elemen penting dalam menentukan posisi dan arah hyperplane optimal[9]. Menentukan hyperplane dinyatakan sebagai $f(w, b) = X_i \cdot w + b$, di mana fungsi tersebut berperan untuk memisahkan data menjadi dua kelompok, yaitu kelas positif dan kelas negatif.

$$X_i \cdot w + b \geq 1 \text{ untuk } y_i = 1 \quad (1)$$

$$X_i \cdot w + b \leq -1 \text{ untuk } y_i = -1 \quad (2)$$

Keterangan:

X_i : data ke i

W : bobot *support vector* atau vector yang tegak lurus terhadap hyperplane

B : nilai bias

Y_1 : kelas data ke i

Untuk menentukan posisi hyperplane optimal, yaitu bidang pemisah yang berada di antara keduanya dengan jarak maksimum dari masing-masing kelas.

$$\text{margin} = |d_{H1} - d_{H2}| = \frac{2}{\|w\|} \quad (3)$$

Keterangan:

d_{H1} : jarak hyperplane pendukung kelas +1

d_{H2} : jarak hyperplane pendukung kelas -1

Pada kondisi data yang memiliki pola nonlinier, metode pemisahan secara linear tidak lagi memadai, sehingga diperlukan pendekatan tambahan melalui penggunaan fungsi kernel. Fungsi ini bekerja dengan mentransformasikan data menjadi mudah dipisahkan. Salah satu kernel yang paling sering digunakan adalah *Radial Basis Function* (RBF), yang memiliki kemampuan baik dalam menangani pola data yang kompleks dan dapat dirumuskan sebagai berikut:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (4)$$

Pada persamaan kernel x_i dan x_j merepresentasikan pasangan data latih, sedangkan parameter $\sigma > 0$ merupakan suatu konstanta yang mengatur tingkat penyebaran fungsi. Kernel berperan sebagai pengganti operasi dot *product* pada ruang fitur, yang sangat dipengaruhi oleh karakteristik data. Metode SVM dikembangkan untuk menangani klasifikasi dua kelas. Namun, seiring perkembangannya, SVM telah dimodifikasi untuk dapat digunakan pada kasus klasifikasi multikelas. Dalam proses ini, akan terbentuk beberapa hyperplane yang digunakan untuk memisahkan masing-masing kelas dalam data.

2.6. Confusion Matrix

Confusion matrix merupakan table yang digunakan untuk menilai sejauh mana model klasifikasi mampu menghasilkan prediksi yang akurat[10]. Model *confusion matrix* dapat dilihat pada gambar berikut:

		Predicted	
		Positive	Negative
Actual	Positive	TP	FP
	Negative	FN	TN

Gambar 1. Confusion Matrix

a. Akurasi

Akurasi merupakan perbandingan antara jumlah data yang berhasil diklasifikasikan dengan benar terhadap seluruh data dalam proses pengujian. Menghitung nilai akurasi menggunakan persamaan berikut:

$$akurasi = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (5)$$

b. Presisi

Presisi merupakan perbandingan antara jumlah data yang diklasifikasikan benar sebagai kategori positif dengan seluruh data yang diprediksi sebagai positif. Menghitung nilai presisi menggunakan persamaan berikut:

$$Presisi = \frac{TP}{TP + FP} \times 100\% \quad (6)$$

c. Recall

Recall merupakan presentase data kategori positif yang berhasil diklasifikasikan dengan benar yang dibandingkan dengan seluruh data positif sebenarnya. Menghitung nilai recall menggunakan persamaan berikut:

$$Recall = \frac{TP}{FP + FN} \times 100\% \quad (7)$$

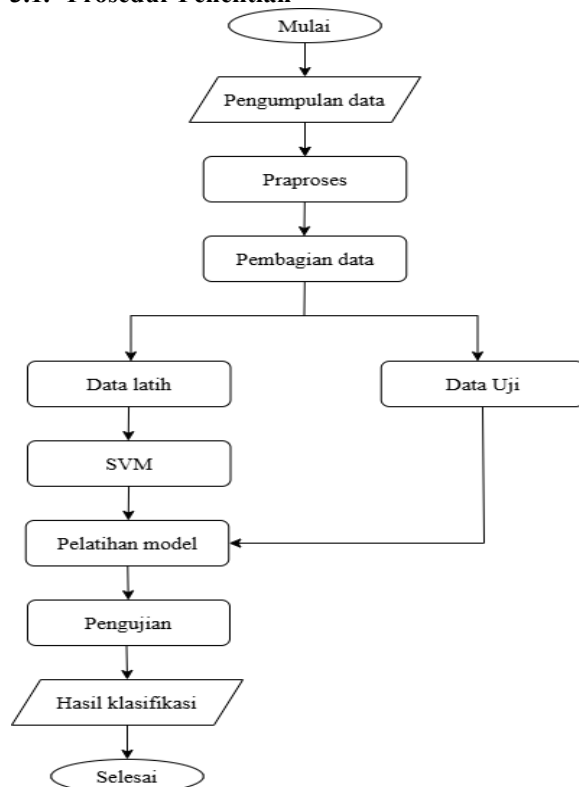
d. F1-Score

F1-score merupakan nilai rata-rata dari presisi dan recall. Untuk menghitung nilai recall menggunakan persamaan berikut:

$$F1 - Score = 2 \times \frac{presisi \times recall}{presisi + recall} \times 100\% \quad (8)$$

3. METODE PENELITIAN

3.1. Prosedur Penelitian



Gambar 2. Flowchart Penelitian

Prosedur penelitian ini menjelaskan tahapan yang digunakan dalam mengolah data menjadi informasi serta menguji model klasifikasi status gizi balita menggunakan algoritma *Support Vector Machine* (SVM). Tahapan tersebut meliputi pengumpulan data, praproses data, normalisasi, pelatihan model hingga evaluasi hasil klasifikasi. Rangkaian prosedur penelitian ini ditampilkan pada gambar 2.

3.2. Data Balita

Data yang digunakan berupa data balita yang diperoleh dari Puskesmas Oebobo Kota Kupang dengan jumlah sebanyak 475. Data tersebut berupa jenis kelamin, usia, berat badan, dan tinggi badan. Distribusi data berdasarkan kategori status gizi menunjukkan bahwa jumlah balita dengan status gizi normal yang paling dominan, yaitu sekitar 282 data, gizi kurang 82 data, gizi buruk 55 data dan 57 data untuk gizi lebih.

3.3. Praproses

Praproses dilakukan dengan mentransformasi data pada atribut jenis kelamin dengan mengubah data kategorikal menjadi numerik. Selanjutnya dilakukan proses normalisasi data agar menyamakan skala antar fitur.

3.4. Pembagian Data

Pembagian data latih dan data uji menggunakan *train_test_split* dengan perbandingan 80% data latih dan 20% data uji.

3.5. Model SVM

Model *Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF) dilatih menggunakan data latih yang telah melalui proses preprocessing. Untuk memperoleh performa optimal, dilakukan pengujian kombinasi parameter C (1, 10, 50) dan gamma (scale, 0.1, 0.001). Setiap kombinasi diuji dengan melatih model dan menghitung akurasi pada data uji. Model dengan akurasi tertinggi dipilih sebagai model terbaik untuk proses klasifikasi.

4. HASIL DAN PEMBAHASAN

Proses klasifikasi model *Support Vector Machine* (SVM) *multiclass* dengan kernel RBF menggunakan variasi parameter C = 1, 10, 50 dan gamma = scale, 0.1, 0.01 untuk memperoleh kombinasi parameter terbaik. Evaluasi model dilakukan menggunakan akurasi, presisi, recall, dan f1-score dengan metode *weighted average* agar hasil evaluasi mempertimbangkan proporsi jumlah data pada setiap kelas, sehingga lebih representatif terhadap kondisi data.

Berikut hasil akurasi pengujian dari perhitungan tersebut di tampilkan pada tabel 1.

Tabel 1. Hasil Akurasi Pengujian

C	Gamma		
	scale	0,1	0,01
1	69,79%	70,80%	59,38%
10	79,17%	78,12%	71,88%
50	80,21%	78,12	73,12

Berdasarkan hasil pengujian variasi parameter C dan gamma pada metode *Support Vector Machine* dengan kernel RBF, terlihat performa model meningkat seiring dengan penggunaan parameter yang lebih optimal. Pada nilai C yang kecil, seperti C = 1,

akurasi masih rendah karena model belum mampu memisahkan data dengan baik. Ketika nilai C ditingkatkan menjadi C = 10, akurasi mulai meningkat karena model lebih tegas dalam menentukan batas antar kelas. Peningkatan performa lebih baik terlihat pada C = 50, terutama pada penggunaan gamma = scale, dimana model mampu mengklasifikasikan data lebih akurat. Namun, pada nilai gamma yang terlalu kecil seperti 0.01, performa model cenderung menurun karena model kurang mampu menangkap pola data secara optimal. Berikut adalah hasil pengujian dengan *confusion matrix* pada nilai C = 50 dan gamma = scale.

Tabel 2. *Confusion Matrix*

		Predicted			
		Gizi Buruk	Gizi Kurang	Gizi Normal	Gizi Lebih
Actual	Gizi Buruk	6	2	3	0
	Gizi Kurang	3	8	6	0
	Gizi Normal	1	2	53	1
	Gizi Lebih	0	0	1	10

Berdasarkan hasil *confusion matrix* dari data balita didapatkan nilai akurasi, presisi, recall dan f1-score sebagai berikut:

a. Akurasi

$$Akurasi = \frac{6 + 8 + 53 + 10}{96} \times 100\% = 80,21\%$$

b. Presisi

$$\begin{aligned} \text{Presisi} &= \left[\left(\frac{6}{6+4} \right) \times 11 \right] + \left[\left(\frac{8}{8+4} \right) \times 17 \right] + \\ &\quad \left[\left(\frac{53}{53+10} \right) \times 57 \right] + \left[\left(\frac{10}{10+1} \right) \times 11 \right] \\ &= 79,05\% \end{aligned}$$

c. Recall

$$\begin{aligned} \text{Recall} &= \left[\left(\frac{6}{6+5} \right) \times 11 \right] + \left[\left(\frac{8}{8+9} \right) \times 17 \right] + \\ &\quad \left[\left(\frac{53}{53+4} \right) \times 57 \right] + \left[\left(\frac{10}{10+1} \right) \times 11 \right] \\ &= 80,21\% \end{aligned}$$

d. F1-Score

$$\begin{aligned} \text{F1-Score} &= \left[\left(2 \times \frac{0,6 \times 0,4545}{0,6 + 0,4545} \right) \times 11 \right] + \\ &\quad \left[\left(2 \times \frac{0,6666 \times 0,4705}{0,6666 + 0,4705} \right) \times 17 \right] + \\ &\quad \left[\left(2 \times \frac{0,8412 \times 0,9228}{0,8412 + 0,9228} \right) \times 57 \right] + \\ &\quad \left[\left(2 \times \frac{0,9090 \times 0,9090}{0,9090 + 0,9090} \right) \times 11 \right] \\ &= 79,18\% \end{aligned}$$

Dari perhitungan *confusion matrix* pada parameter C = 50 dan gamma = scale didapatkan hasil akurasi 80,21%, presisi 79,05%, dan f1-score 79,18%.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa model SVM dengan kernel RBF berhasil digunakan untuk mengklasifikasikan status gizi balita dengan hasil pengujian menunjukkan bahwa nilai performa model dipengaruhi oleh pemilihan parameter C dan gamma, di mana nilai parameter terbaik diperoleh pada kombinasi C = 50 gamma = scale dengan akurasi 80,21%. Saran dari penelitian ini, menggunakan jumlah data yang lebih besar dan distribusi kelas yang lebih seimbang agar performa model lebih meningkat.

DAFTAR PUSTAKA

- [1] A. Nadroh, D. N. Triwibowo, and B. S. R. Bagus, "Klasifikasi Status Gizi Balita Menggunakan Algoritma Support Vector Machine dengan Optimasi Grid Search Cross-Validation," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 8, no. 2, pp. 250–257, 2024, doi: 10.46880/jmika.vol8no2.pp250-257.
- [2] E. Ramon, A. Nazir, N. Novriyanto, Y. Yusra, and L. Oktavia, "Classification of Infant Nutritional Status at Bangun Purba Subdistrict Integrated Health Post Using Support Vector Machine (SVM) Algorithm," *J. Sist. Inf. dan Inform.*, vol. 5, no. 2, pp. 143–150, 2022, [Online]. Available: <https://ejournal.lppm-unbaja.ac.id/index.php/jsii/article/view/2185>
- [3] Kementerian Kesehatan, "Survei Status Gizi Indonesia 2024," 2024, [Online]. Available: <https://www.badankebijakan.kemkes.go.id/survei-status-gizi-indonesia-ssgi-2024/>
- [4] J. T. Terapan, "Jurnal Teknologi Terapan & Sains 4.0," vol. 5, no. 2, 2024.
- [5] A. Candra, M. Erkamim, M. Muharrom, and E. Prayitno, "Klasifikasi Stunting Pada Balita Berdasarkan Status Gizi Menggunakan Pendekatan Support Vector Machine (SVM)," *J. Ilm. FIFO*, vol. 16, no. 2, p. 171, 2024, doi: 10.22441/fifo.2024.v16i2.007.
- [6] M. Yoshe and W. Hadikurniawati, "Implementasi Metode Naive Bayes Classifier Untuk Klasifikasi Status Gizi Stunting Pada Balita," vol. 2019, 2021.
- [7] P. Handayani and A. Charis Fauzan, "Machine Learning Klasifikasi Status Gizi Balita Menggunakan Algoritma Random Forest," *Media Online*, vol. 4, no. 6, pp. 3064–3072, 2024, doi: 10.30865/klik.v4i6.1909.
- [8] S. Rahayu and Y. Yamasari, "Klasifikasi Penyakit Stroke dengan Metode Support Vector Machine (SVM)," vol. 05, pp. 440–446, 2024.
- [9] R. Nanda, E. Haerani, S. K. Gusti, and S. Ramadhani, "Klasifikasi Berita Menggunakan Metode Support Vector Machine," vol. 5, no. 2, pp. 269–278, 2022.
- [10] G. Rininda, I. H. Santi, and S. Kirom, "Penerapan Svm Dalam Analisis Sentimen Pada Edlink," vol. 7, no. 5, pp. 3335–3342, 2023.