



# Studi Perbandingan Performa: Prediksi Status *Stunting* Pada Anak Berdasarkan Data Antropometri Menggunakan Algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN)

Gita Rahmawati<sup>1</sup>, Safril Ahmadi Sanmas<sup>2</sup>, Elsa Nudyawati<sup>3</sup>, Nabbila Dyah Syaharani<sup>4</sup>, Mochamad Raffy Annas<sup>5</sup>, Fatkhurokhan Fauzi<sup>6</sup>

<sup>1,2,3,4,5,6</sup>SI Statistika, Universitas Muhammadiyah Semarang

<sup>1</sup>[gitarahmawati01@gmail.com](mailto:gitarahmawati01@gmail.com), <sup>2</sup>[safriksanmaso02@gmail.com](mailto:safriksanmaso02@gmail.com), <sup>3</sup>[elsanudyawati2001@gmail.com](mailto:elsanudyawati2001@gmail.com), <sup>4</sup>[nabbilads@gmail.com](mailto:nabbilads@gmail.com),

<sup>5</sup>[raffyan1204@gmail.com](mailto:raffyan1204@gmail.com), <sup>6</sup>[fatkhurokhanff@unimus.ac.id](mailto:fatkhurokhanff@unimus.ac.id)

Corresponding author email: [gitarahmawati01@gmail.com](mailto:gitarahmawati01@gmail.com)

**Abstract:** *Stunting is a condition of impaired growth due to long-term malnutrition, threatening the physical health and cognitive development of children. The Indonesian Health Survey (SKI) 2023 revealed that the stunting prevalence in Indonesia reached 21.5%, far from the target of 14% set for 2024. Determining stunting status manually through weight and height measurements has limitations, such as dependency on staff skills and inefficiency for large populations. Therefore, machine learning technologies like Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) offer innovative solutions with high accuracy. SVM searches for the best hyperplane to separate stunting and non-stunting data, while KNN classifies new data based on its proximity to existing data. This study aims to compare the performance of SVM and KNN in predicting stunting status. The results showed that SVM using an RBF kernel and parameter  $C=10$  achieved an accuracy of 97.11%, precision of 97.10%, recall of 97.11%, and F1-score of 97.10%. Meanwhile, the KNN method with  $k=1$  achieved an accuracy of 97.05%, precision of 97.05%, recall of 97.05%, and F1-score of 97.05%. Therefore, SVM is superior in accuracy and efficiency for diagnosing stunting, providing a more precise, faster, and broader solution to address the stunting problem.*

**Keywords:** *Stunting, SVM, KNN*

**Abstrak:** *Stunting adalah kondisi gagal tumbuh akibat kekurangan gizi jangka panjang, mengancam kesehatan fisik dan perkembangan kognitif anak. Survei Kesehatan Indonesia (SKI) 2023 mengungkap data prevalensi stunting di Indonesia mencapai 21,5%, jauh dari target 14% yang ditargetkan untuk tahun 2024. Penentuan status stunting secara manual melalui pengukuran berat dan tinggi badan memiliki keterbatasan, seperti ketergantungan pada keterampilan petugas dan kurang efisien untuk populasi besar. Oleh karena itu, teknologi *machine learning*, seperti *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN), menawarkan solusi inovatif dengan akurasi tinggi. SVM mencari *hyperplane* terbaik untuk memisahkan data *stunting* dan *non-stunting*, sementara KNN mengklasifikasikan data baru berdasarkan kedekatan dengan data yang sudah ada. Penelitian ini bertujuan untuk membandingkan performa SVM dan KNN dalam memprediksi status *stunting*. Hasil penelitian menunjukkan bahwa SVM dengan menggunakan kernel RBF dan parameter  $C=10$  mencapai akurasi sebesar 97,11%, presisi 97,10%, *recall* 97,11%, dan *F1-score* 97,10%. Sementara itu, metode KNN dengan  $k=1$  mencapai akurasi 97,05%, presisi 97,05%, *recall* 97,05%, dan *F1-score* 97,05%. Dengan demikian, SVM lebih unggul dalam akurasi dan efisiensi untuk diagnosis *stunting* dengan memberikan solusi yang lebih tepat, cepat, dan luas dalam menangani permasalahan *stunting*.*

**Kata kunci:** *Stunting, SVM, KNN*

## I. PENDAHULUAN

*Stunting* adalah kondisi gagal tumbuh yang diakibatkan oleh asupan gizi tidak memadai dalam jangka panjang, menjadi ancaman serius bagi masa depan anak. *Stunting* tidak hanya berdampak pada kesehatan fisik, seperti peningkatan angka kematian dan kerentanan terhadap penyakit, tetapi juga perkembangan kognitif dan kecerdasan anak [1], [2]. *Stunting* umumnya terjadi pada balita usia 1-3 tahun dan dipengaruhi oleh berbagai faktor, termasuk gizi buruk ibu saat hamil, anemia pada ibu hamil, pola asuh yang tidak tepat, kemiskinan, sanitasi yang buruk, dan pengetahuan gizi yang rendah [3].



Kondisi ini membahayakan potensi anak untuk mencapai kemampuan maksimal dan berakibat buruk bagi masa depan bangsa.

Survei Kesehatan Indonesia (SKI) 2023 mengungkap data prevalensi *stunting* di Indonesia mencapai 21,5%, jauh dari target 14% yang ditargetkan untuk tahun 2024 [4]. Hal ini menunjukkan masih tingginya angka anak yang mengalami gagal tumbuh akibat kekurangan gizi kronis. Saat ini penentuan status *stunting* masih dilakukan secara manual dengan mengukur berat badan dan tinggi badan. Cara ini memiliki keterbatasan diantaranya ketergantungan pada keterampilan petugas dan kurang efisien dalam segi waktu dan tenaga petugas dengan populasi yang besar. Keterbatasan metode manual ini menjadi salah satu hambatan dalam upaya penanggulangan *stunting* di Indonesia [5].

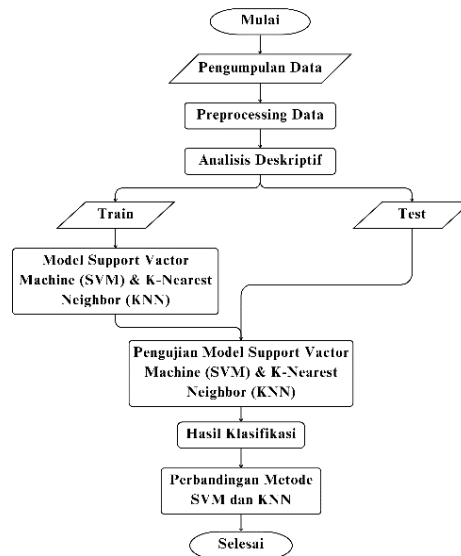
Di era digital ini, teknologi *machine learning* menawarkan solusi inovatif untuk mengatasi permasalahan *stunting*. Algoritma seperti *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN) mampu menganalisis data antropometri dan memprediksi status *stunting* dengan akurasi tinggi. SVM bekerja dengan mencari *hyperplane* terbaik yang memisahkan data anak *stunting* dan *non-stunting*. *Hyperplane* ini memaksimalkan margin antara kedua kelas data. SVM memiliki performa yang baik dalam menangani data yang kompleks dan *non-linear*. KNN, di sisi lain, bekerja dengan membandingkan data baru dengan data yang sudah ada, dan mengklasifikasikannya berdasarkan kedekatan dengan data yang sudah dikategorikan sebelumnya. KNN mudah diimplementasikan dan dipahami, dan memiliki performa yang baik untuk data yang memiliki pola yang jelas [6].

Penelitian mengenai SVM dan KNN telah banyak dilakukan sebelumnya. Penelitian SVM oleh Tineges, Triayudi, dan Sholihati (2020) yang menganalisis sentimen terhadap layanan Indihome berdasarkan Twitter menunjukkan hasil akurasi 87%, presisi 86%, *recall* 95%, dan *F1-score* 90% [7]. Selain itu penelitian mengenai KNN oleh Putra, Pardede, dan Syahputra (2022) menganalisis klasifikasi data iris bunga menunjukkan hasil akurasi sebesar 100% pada semua nilai K yang diuji [8].

Penelitian ini berfokus pada pengembangan model prediksi status *stunting* pada anak yang lebih akurat dan efisien dibandingkan metode manual yang biasa digunakan. Kedua algoritma *machine learning*, KNN dan SVM, akan diuji untuk melihat kemampuannya dalam memprediksi status *stunting*. Tujuan dari penelitian ini adalah untuk membandingkan performa KNN dan SVM untuk mengidentifikasi algoritma yang dapat memberikan prediksi status *stunting* dengan akurasi tertinggi. Selain itu, penelitian ini bertujuan untuk memberikan solusi yang lebih tepat dalam mendiagnosis *stunting*, dengan meminimalisir kesalahan dan inkonsistensi yang mungkin terjadi pada metode manual. Dengan demikian, diharapkan penelitian ini dapat mempercepat proses diagnosis *stunting*, menghemat waktu dan tenaga, serta memungkinkan jangkauan yang lebih luas.

## II. METODE PENELITIAN

Metodologi penelitian ini mencakup beberapa tahapan, yaitu pengumpulan data, *preprocessing* data, pemodelan, dan evaluasi model. Gambaran penelitian dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

### 2.1 Pengumpulan Data

Penelitian ini menggunakan *Dataset Stunting Baby/Toddler Detection* yang terdiri dari 10.000 baris dan 4 kolom. Variabel yang digunakan adalah umur (*age*), jenis kelamin (*gender*), panjang badan (*body length*), dan status gizi (*nutritional status*) sebagai label. Analisis variabel dilakukan terhadap jumlah data untuk setiap atribut yang dijelaskan pada tabel 1.

Tabel 1. Data Stunting

Umur (bulan)	Jenis Kelamin	Tinggi Badan (cm)	Status Gizi
0	laki-laki	44.59	stunted
0	laki-laki	56.71	tinggi
0	laki-laki	46.86	normal
0	laki-laki	47.51	normal
0	laki-laki	42.74	severely
.	.	.	.
60	perempuan	116.1	normal
60	perempuan	100.6	normal
60	perempuan	98.3	stunted
60	perempuan	121.3	normal
60	perempuan	112.2	normal
60	perempuan	109.8	normal

Dimana:

*Stunted*: menunjukkan kondisi *stunting*.

Tinggi: menunjukkan pertumbuhan di atas rata-rata.

Normal: menunjukkan status gizi yang sehat.

*Severely*: menunjukkan kondisi *stunting* yang sangat serius.



## 2.2 Data Preprocessing

*Preprocessing* adalah tahap persiapan sebelum analisis data yang sebenarnya dimulai. Pada tahap ini, dilakukan pembersihan dan pemrosesan data, integrasi serta pengurangan data, transmisi data, dan normalisasi data [9].

## 2.3 Data Classification

Salah satu tugas utama dalam *data mining* adalah klasifikasi, yang termasuk dalam pembelajaran terawasi. Klasifikasi adalah proses membuat model atau fungsi yang menggambarkan atau membedakan suatu konsep atau kelas data, dengan tujuan mengidentifikasi kelas objek yang labelnya belum diketahui. Untuk mencapai tujuan ini, model dibangun melalui proses kategorisasi yang mengelompokkan data ke dalam beberapa kelas berdasarkan aturan atau fungsi yang telah ditetapkan. Kinerja klasifikasi diukur menggunakan *confusion matrix* [10].

## 2.4 K-Nearest Neighbor

*K-Nearest Neighbor* (K-NN) adalah salah satu algoritma *Machine Learning* yang digunakan untuk tugas klasifikasi dan regresi. K-NN adalah algoritma yang sederhana namun efektif. KNN mengklasifikasikan objek atau contoh data berdasarkan kelas mayoritas dari tetangga terdekatnya. Algoritma ini termasuk dalam kategori pembelajaran berbasis instan (*instance-based learning*). Selain itu, *K-Nearest Neighbor* memiliki keunggulan dalam hal kemudahan implementasi dan kinerja yang baik dalam menangani dataset yang besar. Perhitungan jarak antara *data training* dan *data testing* menggunakan persamaan (1) [6].

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2} \quad (1)$$

keterangan :

$x_{1i}$  : Data latih atau *training* dari dataset yang akan diklasifikasi

$x_{2i}$  : Data uji atau *testing* dari dataset yang akan diklasifikasi

## 2.5 Support Vector Machine

*Support Vector Machine* (SVM) adalah algoritma *machine learning* yang digunakan untuk klasifikasi dan regresi dengan mencari garis atau permukaan pemisah optimal antara kelas-kelas data. SVM unggul dalam menangani ruang dimensi tinggi dan data tidak linear, serta dapat diterapkan pada masalah klasifikasi multi-kelas. Dalam konteks klasifikasi, SVM bertujuan menemukan *hyperplane* terbaik untuk memisahkan dua kelas di ruang input dengan margin sebagai jarak antara pola terdekat dari setiap kelas dan *hyperplane*. Vektor pendukung adalah pola-pola yang paling dekat dengan *hyperplane*. Fungsi keputusan SVM dapat dilihat dengan persamaan (2) [6].

$$f(x_d) = \sum_{i=1}^{n_s} a_i y_i x_i x_d + b \quad (2)$$

Dimana:

$N_s$  : Jumlah *support vector*

$a_i$  : Besaran bobot untuk setiap titik data

$y_i$  : Kelas data

$x_i$  : variabel *support vector*

$x_d$  : Data yang akan diklasifikasikan

$b$  : Nilai *error*



## 2.6 Evaluasi Model

Untuk mengetahui hasil kinerja klasifikasi dari model algoritma, nilai *Confusion Matrix* dapat digunakan. *Confusion Matrix* ini akan merepresentasikan kebenaran dari sebuah prediksi dengan memperlihatkan jumlah *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN). Dengan *Confusion Matrix*, dapat dihitung hitung metrik evaluasi akurasi, presisi, *recall*, dan *F1-score* yang memberikan gambaran lengkap tentang performa model klasifikasi [11]. Persamaan fungsi kinerja masing-masing metrik evaluasi dapat dilihat pada tabel 2 [12].

**Tabel 2.** Rumus Fungsi Kinerja Klasifikasi

Ukuran Kinerja	Rumus
Akurasi	$\frac{TP + TN}{TP + TN + FP + FN}$
Presisi	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1-Score	$\frac{2(Presisi \times Recall)}{(Presisi + Recall)}$

## III. HASIL DAN PEMBAHASAN

### 3.1. Pembersihan Data

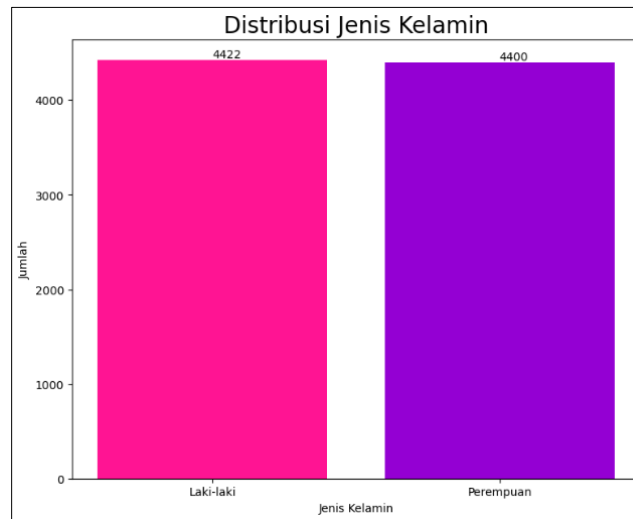
Penelitian ini menggunakan modul *pandas* pada *python* sebagai alat untuk pembersihan data. Setelah melakukan pembersihan data, jumlah data menjadi 8.822 data dikarenakan terdapat data duplikat sebanyak 1.178 data. Penghapusan data duplikat penting dilakukan untuk menghindari bias dalam model prediksi dan memastikan bahwa setiap entri data unik dan representatif dari populasi sebenarnya [13].

### 3.2. Analisis Deskriptif

**Tabel 3.** Deskriptif Data

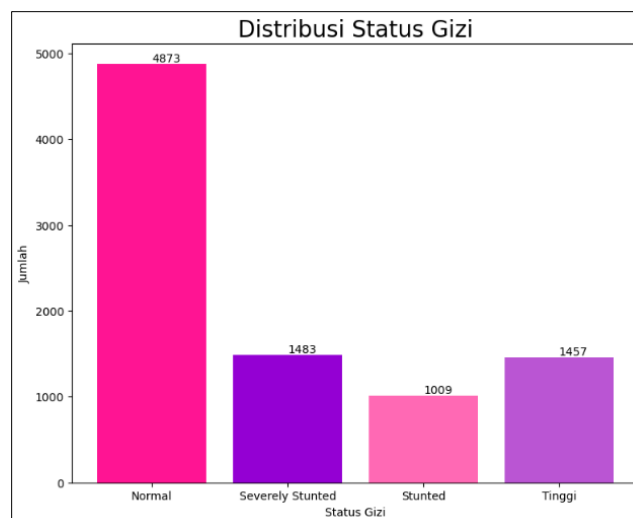
	Min	Mean	Max	Std
Umur (Bulan)	0,0000	29,6123	60,0000	17,5949
Tinggi Badan (cm)	40,0104	88,1103	127,8000	17,4436

Berdasarkan tabel 3, umur anak-anak dalam dataset berkisar dari 0 hingga 60 bulan dengan rata-rata 29,6 bulan dan simpangan baku 17,6 bulan. Tinggi badan berkisar dari 40,01 cm hingga 128 cm dengan rata-rata 88,1 cm dan simpangan baku 17,4 cm. Data ini menunjukkan adanya variasi signifikan dalam umur dan tinggi badan anak-anak yang penting untuk memahami perkembangan mereka dan dapat menjadi dasar untuk penelitian lebih lanjut atau intervensi yang diperlukan.



Gambar 2. Distribusi Jenis Kelamin

Berdasarkan gambar 2, terdapat 4.422 anak laki-laki dan 4.400 anak perempuan. Data ini menunjukkan bahwa jumlah anak laki-laki sedikit lebih banyak dibandingkan anak perempuan dalam dataset tersebut.



Gambar 3. Distribusi Status Gizi

Berdasarkan gambar 3, terdapat 4.873 anak termasuk dalam kategori "normal," 1.483 anak "Severely Stunted", 1.009 anak "Stunted", dan 1.457 anak "tinggi". Mayoritas anak berada dalam kategori "normal", namun terdapat juga proporsi yang signifikan dari anak-anak yang mengalami masalah gizi, baik dalam bentuk kekurangan maupun kelebihan gizi.

### 3.3. Support Vector Machine

Support Vector Machine (SVM) digunakan untuk menganalisis *dataset* yang terdiri dari total 8.822 data, yang dibagi menjadi 80% untuk data *training* dan 20% untuk data *testing*. Penggunaan SVM dilengkapi dengan kernel RBF yang merupakan fungsi kernel untuk menganalisis ketika data tidak dapat dipisahkan secara linear. RBF kernel memiliki dua parameter utama yaitu parameter Gamma ( $\gamma$ ) dan Cost (C). Parameter terbaik yang ditemukan adalah  $C=10$  dan  $\gamma='scale'$ . Parameter dipilih



setelah serangkaian eksperimen untuk mencapai keseimbangan antara akurasi dan generalisasi model [14].

Untuk menghindari kesalahan prediksi, validasi dilakukan menggunakan metode *cross validation* dengan pembagian data *training* menjadi 5 *fold*. Proses ini memastikan bahwa parameter yang dipilih tidak hanya optimal untuk data *training* tetapi juga dapat menggeneralisasi dengan baik pada data *test* [15]. Setelah menentukan parameter terbaik melalui *Grid Search*, model SVM yang optimal diuji menggunakan data testing. Hasil prediksi model SVM dengan parameter terbaik menghasilkan *confusion matrix* yang dapat dilihat pada tabel 4.

**Tabel 4.** Confusion Matrix SVM

n = 1767	Predicted 0	Predicted 1	Predicted 2	Predicted 3
Actual 0	926	1	5	10
Actual 1	2	324	7	0
Actual 2	11	7	179	0
Actual 3	10	0	0	285

Berdasarkan tabel 4, dapat dilihat bahwa status gizi yang diklasifikasikan dengan benar pada kategori 0 (normal) sebanyak 926, kategori 1 (*Severely Stunted*) sebanyak 324, kategori 2 (*Stunted*) sebanyak 179, dan kategori 3 (tinggi) sebanyak 285.

### 3.4. K-Nearest Neighbor

*K-Nearest Neighbor* (KNN) digunakan untuk menganalisis dataset yang terdiri dari total 8.822 data, yang dibagi menjadi 80% untuk data *training* dan 20% untuk data *testing*. Pemilihan nilai K yang optimal dilakukan melalui proses *cross-validation* untuk memastikan bahwa model dapat menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya. Dalam penelitian ini, nilai K yang terpilih adalah  $K = 1$ . Hasil prediksi model KNN dengan parameter terbaik menghasilkan *confusion matrix* yang dapat dilihat pada tabel 5.

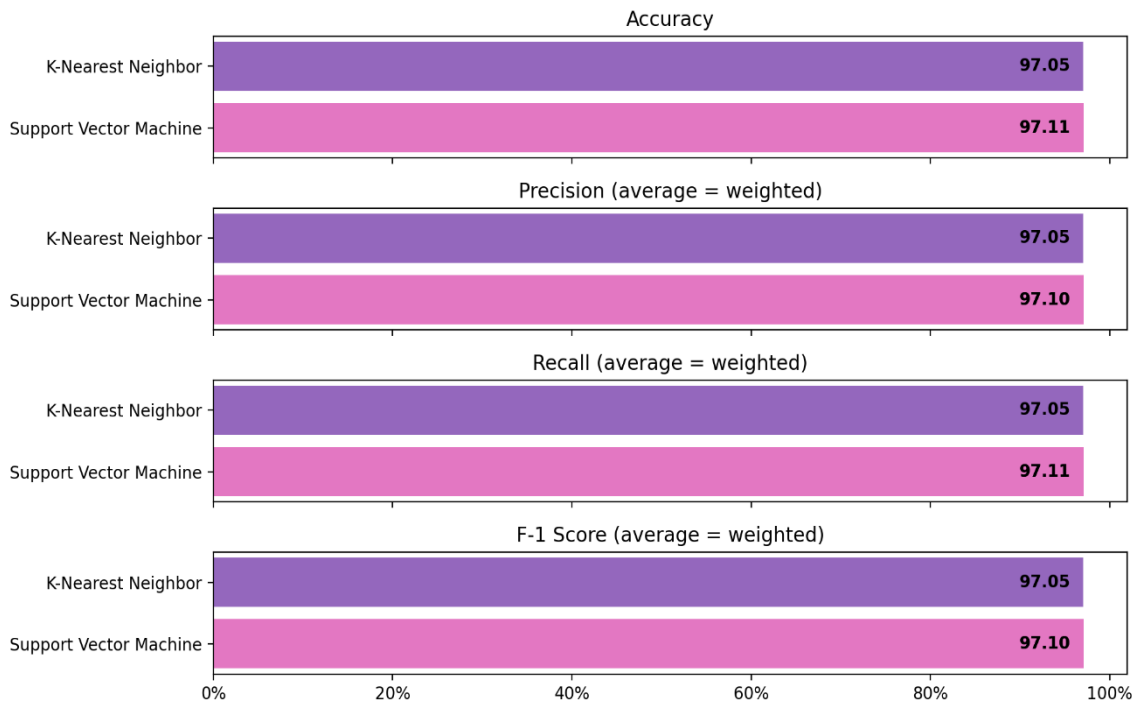
**Tabel 5.** Confusion Matrix K = 1

n = 1767	Predicted 0	Predicted 1	Predicted 2	Predicted 3
Actual 0	922	0	8	12
Actual 1	0	324	7	0
Actual 2	9	7	181	0
Actual 3	9	0	0	286

Berdasarkan tabel 5, dapat dilihat bahwa status gizi yang diklasifikasikan dengan benar pada kategori 0 (normal) sebanyak 922, kategori 1 (*Severely Stunted*) sebanyak 324, kategori 2 (*Stunted*) sebanyak 181, dan kategori 3 (tinggi) sebanyak 286.

### 3.4. Evaluasi Model

Evaluasi model dilakukan dengan menggunakan akurasi, presisi, *recall*, dan *F1-score*. Hasil evaluasi model dapat dilihat pada gambar 4.



Gambar 4. Perbandingan Model

Berdasarkan gambar 4, diperoleh hasil bahwa model SVM memiliki akurasi sebesar 97,11%, presisi 97,10%, *recall* 97,11%, dan *F1-score* 97,10%. Sementara model KNN memiliki akurasi 97,05%, presisi 97,05%, *recall* 97,05%, dan *F1-score* 97,05%. Secara keseluruhan, terlihat bahwa model SVM menunjukkan kinerja yang lebih baik dibandingkan dengan model KNN.

#### IV. KESIMPULAN

Berdasarkan analisis yang dilakukan menggunakan metode SVM dan KNN untuk memprediksi status gizi anak, hasil menunjukkan bahwa SVM dengan kernel RBF dan parameter  $C=10$  mencapai akurasi 97,11%, presisi 97,10%, *recall* 97,11%, dan *F1-score* 97,10%. Sedangkan, metode KNN dengan  $K=1$  mencapai akurasi 97,05%, presisi 97,05%, *recall* 97,05%, dan *F1-score* 97,05%. Meskipun kedua metode menunjukkan performa yang tinggi dan hampir seimbang, SVM sedikit lebih unggul dalam hal akurasi dan efisiensi. Dengan demikian, dalam konteks prediksi status gizi anak, terutama untuk diagnosis stunting, metode SVM dengan kernel RBF dan parameter yang dioptimalkan menjadi pilihan yang lebih baik. Hasil ini menunjukkan bahwa SVM dapat memberikan kontribusi yang signifikan dalam menangani permasalahan stunting secara lebih luas dan efektif.

#### REFERENSI

1. Fitriani *et al.*, “Cegah Stunting Itu Penting!,” *Jurnal Pengabdian Kepada Masyarakat (JurDikMas) Sosiosaintifik*, vol. 4, no. 2, pp. 63–67, 2022, doi: 10.54339/jurdikmas.v4i2.417.
2. S. Syahrial, R. Ilham, Z. F. Asikin, and St. S. I. Nurdin, “Stunting Classification in Children’s Measurement Data Using Machine Learning Models,” *Journal La Multiapp*, vol. 3, no. 2, pp. 52–60, 2022, doi: 10.37899/journallamultiapp.v3i2.614.
3. K. Komalasari, E. Supriati, R. Sanjaya, and H. Ifayanti, “Faktor-Faktor Penyebab Kejadian Stunting Pada Balita,” *Majalah Kesehatan Indonesia*, vol. 1, no. 2, pp. 51–56, Oct. 2020, doi: 10.47679/makein.202010.





4. stunting.go.id, “Stunting - Tim Percepatan Pencegahan Anak Kerdil,” *Website Kementerian sekretariat negara RI*. 2021. Accessed: Jun. 10, 2024. [Online]. Available: <https://stunting.go.id/>
5. M. Ludya, Y. Herlambang, and D. Yunidar, “Produk alat ukur tinggi dan berat badan pendeteksi stunting dengan fitur hiburan untuk anak usia 2-5 tahun,” *Productum: Jurnal Desain Produk (Pengetahuan dan Perancangan Produk)*, vol. 6, no. 1, pp. 51–62, 2023, doi: 10.24821/productum.v6i1.7685.
6. Shedriko, “Perbandingan Algoritma SVM dan KNN dalam Mengklasifikasikan Kelulusan Mahasiswa Pada Suatu Mata Kuliah,” vol. 6, no. 2, pp. 115–122, 2021.
7. R. Tineges, A. Triayudi, and I. D. Sholihati, “Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM),” *Jurnal Media Informatika Budidarma*, vol. 4, no. 3, p. 650, Jul. 2020, doi: 10.30865/mib.v4i3.2181.
8. P. Putra, A. M. H. Pardede, and S. Syahputra, “Analisis Metode K-Nearest Neighbour (KNN) Dalam Klasifikasi Data Iris Bunga,” *Jurnal Teknik Informatika Kaputama (JTik)*, vol. 6, no. 1, pp. 297–305, 2022, Accessed: Jun. 17, 2024. [Online]. Available: <https://jurnal.kaputama.ac.id/index.php/JTik/article/view/401/352>
9. Binus Student Corner, “Teknik pre-processing dan classification dalam data science – Master of Industrial Engineering,” Binus University. Accessed: Jun. 10, 2024. [Online]. Available: <https://mie.binus.ac.id/2022/08/26/teknik-pre-processing-dan-classification-dalam-data-science/>
10. Albert Verasius Dian Sano, “Cara Kerja Data Mining – Seri Data Mining for Business Intelligence (3),” Binus University. Accessed: Jun. 10, 2024. [Online]. Available: <https://binus.ac.id/malang/2019/01/cara-kerja-data-mining-seri-data-mining-for-business-intelligence-3/>
11. M. D. Irfan, *Prediksi Cacat Perangkat Lunak Menggunakan Algoritma Max-Min Ant System dan Teknik Bagging*. Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta, 2023. Accessed: Jun. 10, 2024. [Online]. Available: <https://repository.uinjkt.ac.id/dspace/handle/123456789/76250>
12. S. Irmeilyana, Herlina Hanum, Yuli Andriani, “Penerapan Metode Support Vector Machine Dalam Klasifikasi Bunga Iris,” *Universitas Nusantara PGRI Kediri*, vol. 01, pp. 1–7, 2017.
13. P. M. S. Madani, T. Rohana, K. A. Baihaqi, and A. Fauzi, “Perbandingan Kinerja Klasifikasi Penyakit Ginjal Menggunakan Algoritma Support Vector Machine (SVM) dan Decision Tree (DT),” *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 1, pp. 74–82-74–82, Jun. 2024, doi: 10.47065/BITS.V6I1.5206.
14. Y. Julianto, D. H. Setiabudi, and S. Rostianingsih, “Analisis Sentimen Ulasan Restoran Menggunakan Metode SVM,” *Jurnal Infra*, vol. 10, no. 1, 2022.
15. W. A. Firmansyach, U. Hayati, and Y. Arie Wijaya, “Analisa Terjadinya Overfitting Dan Underfitting Pada Algoritma Naive Bayes Dan Decision Tree Dengan Teknik Cross Validation,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 262–269, 2023, doi: 10.36040/jati.v7i1.6329.