

Jurnal JTik (Jurnal Teknologi Informasi dan Komunikasi)

DOI: <https://doi.org/10.35870/jtik.v9i2.3436>

Pendekatan *Machine Learning* untuk Deteksi *Stunting* pada Balita Menggunakan *K-Nearest Neighbors*

Ade Putra Tupu Djoru ^{1*}, Sri Yulianto ²^{1*,2} Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Kota Salatiga, Provinsi Jawa Tengah, Indonesia.

article info

Article history:

Received 5 December 2024

Received in revised form

20 December 2024

Accepted 1 January 2025

Available online April 2025.

Keywords:

Stunting; K-Nearest

Neighbors; Toddlers;

Anthropometrics.

Kata Kunci:

Stunting; K-Nearest

Neighbors; Balita;

Antropometri.

abstract

Stunting is a chronic nutritional problem affecting the physical growth and cognitive development of toddlers, especially in early childhood. This study employs the K-Nearest Neighbors (K-NN) algorithm to determine stunting status based on anthropometric variables such as age, weight, and height. The algorithm categorizes data using proximity between samples. Data from the Salatiga City Health Department in 2024 were normalized and encoded for analysis. K-NN was chosen for its ability to provide high-accuracy nutritional classification. Results show that the algorithm achieved 100% accuracy, precision, and recall at certain K values, particularly in the small K range (2-8), demonstrating its effectiveness in identifying nutritional status in toddlers. This study is expected to serve as a reference for utilizing technology and data analysis in early stunting detection, aiding healthcare professionals in designing more targeted and effective interventions. Additionally, it opens opportunities for further development to enhance diagnostic accuracy using machine learning technology.

abstrak

Stunting adalah masalah gizi kronis yang memengaruhi pertumbuhan fisik dan perkembangan kognitif balita, terutama pada masa awal kehidupan. Penelitian ini menggunakan algoritma K-Nearest Neighbors (K-NN) untuk menentukan status stunting berdasarkan variabel antropometri seperti usia, berat badan, dan tinggi badan. Algoritma ini mampu mengkategorikan data dengan memanfaatkan kedekatan jarak antar sampel. Data dari Dinas Kesehatan Kota Salatiga tahun 2024 telah dinormalisasi dan diencoding untuk analisis. Algoritma K-NN dipilih karena kemampuannya memberikan klasifikasi status gizi dengan akurasi tinggi. Hasil menunjukkan algoritma ini mencapai akurasi, presisi, dan recall 100% pada beberapa nilai K, terutama dalam rentang K kecil (2-8), membuktikan efektivitasnya dalam mengidentifikasi status gizi balita. Penelitian ini diharapkan dapat menjadi referensi dalam penggunaan teknologi dan analisis data untuk deteksi dini stunting, yang akan membantu tenaga kesehatan dalam merancang intervensi yang lebih tepat sasaran dan efektif. Selain itu, penelitian ini membuka peluang pengembangan lebih lanjut untuk meningkatkan ketepatan diagnosa melalui teknologi machine learning.

Corresponding Author. Email: adedjoru16@gmail.com ^{1}.

Copyright 2025 by the authors of this article. Published by Lembaga Otonom Lembaga Informasi dan Riset Indonesia (KITA INFO dan RISET). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. 



ACM Computing Classification System (CCS)

EBSCOhost

Communication and Mass Media Complete (CMMC)

1. Pendahuluan

Stunting pada balita merupakan masalah gizi kronis yang terjadi secara global. Anak-anak yang mengalami keterlambatan perkembangan seringkali menunjukkan tingkat kecerdasan yang lebih rendah, gangguan berbicara, serta kesulitan dalam menangkap pembelajaran dengan metode yang lebih kompleks (Prasetya *et al.*, 2020). Faktor penyebab *stunting* dapat berasal dari berbagai tahapan, termasuk masa kehamilan, persalinan, menyusui, serta masa nifas. Salah satu penyebabnya adalah pemberian Makanan Pendamping ASI (MPASI) yang tidak mencukupi kebutuhan gizi balita (Mentari & Hermansyah, n.d.). Peran orang tua sangat penting dalam penanganan masalah gizi, terutama *stunting*. Dengan pemahaman yang baik mengenai tanda-tanda, pencegahan, dan dampak *stunting*, orang tua dapat mengambil langkah yang tepat untuk mencegah terjadinya *stunting* (Rahmawati *et al.*, 2019). Perubahan status gizi dipengaruhi oleh berbagai faktor, termasuk peningkatan tingkat pendidikan orang tua, perbaikan sanitasi yang memadai, akses terhadap air bersih, peningkatan kesejahteraan keluarga, serta asupan gizi yang lebih baik untuk ibu dan anak. Selain itu, pemahaman yang lebih baik mengenai keterlibatan dalam program gizi dan pola pengasuhan juga turut mendukung perbaikan status gizi anak.

Di berbagai negara berkembang, penelitian menunjukkan bahwa *stunting* berdampak negatif jangka panjang, mempengaruhi anak dan masa depannya. Anak-anak yang mengalami *stunting* cenderung memiliki pencapaian yang lebih rendah dalam pendidikan, penghasilan yang lebih minim, serta risiko yang lebih tinggi untuk terjebak dalam kemiskinan, yang berpengaruh terhadap kesejahteraan mereka di masa depan. Selain itu, anak-anak yang mengalami *stunting* lebih rentan terhadap obesitas dan penyakit kronis saat dewasa karena postur tubuh yang lebih pendek. Pada perempuan, *stunting* juga dapat meningkatkan risiko melahirkan bayi dengan berat badan lahir rendah dan memperbesar kemungkinan terjadinya komplikasi persalinan. Menurut Bank Dunia, penurunan produktivitas ekonomi sebesar 1,4% di negara-negara berkembang di Asia dan Afrika berkaitan dengan penurunan tinggi badan anak sebesar 1%

secara nasional. Indonesia masih menghadapi tantangan terkait gizi kurang pada balita. Hingga saat ini, pemerintah Indonesia terus berupaya mengurangi prevalensi *stunting* melalui dua pendekatan utama: intervensi sensitif dan intervensi spesifik. Intervensi spesifik bertujuan untuk mengatasi dan mencegah *stunting* secara langsung pada balita dan ibu hamil, melalui pemberian suplemen makanan, zat besi, imunisasi, serta suplementasi mikronutrien seperti zat besi, seng, dan vitamin. Di sisi lain, intervensi sensitif merupakan pendekatan lintas sektor yang menangani faktor sosial ekonomi yang meningkatkan risiko *stunting*, seperti akses terhadap air bersih dan sanitasi, kesehatan mental, bantuan sosial, serta ketahanan pangan (Ridho Al Izzati & Elza Elmira, 2019). Kementerian Kesehatan mengumumkan hasil Survei Status Gizi Indonesia (SSGI) dalam Rapat Kerja Nasional BKKBN, yang menunjukkan bahwa prevalensi *stunting* di Indonesia menurun dari 24,4% pada tahun 2021 menjadi 21,6% pada tahun 2022 (Kementerian Kesehatan Republik Indonesia, 2023). Oleh karena itu, diperlukan tindakan cepat di tingkat nasional, kabupaten, kecamatan, dan desa untuk menetapkan area prioritas intervensi.

Dalam RPJMN, pemerintah menargetkan penurunan angka *stunting* menjadi 14 persen pada tahun 2024. Para kepala daerah harus menangani masalah *stunting* dengan serius. Data yang akurat mengenai tingkat prevalensi gizi dan wilayah prioritas sangat penting dalam membantu pengambil keputusan mengalokasikan sumber daya dan anggaran dengan tepat. Informasi yang akurat dan dapat dipercaya sama pentingnya dengan oksigen dalam menjaga kesehatan kita. Banyak penelitian sebelumnya yang membahas tentang *stunting*, di antaranya yang menyatakan bahwa *stunting* dapat terjadi pada bayi dengan berat badan lahir rendah (BBLR) yang tidak menerima pola pengasuhan yang memadai, sehingga asupan energi dan nutrisinya kurang. Selain itu, ibu hamil yang mengalami kekurangan energi kronis (KEK) atau anemia selama masa kehamilan juga bisa menjadi faktor penyebab *stunting* pada balita (Saeful Bachri *et al.*, 2021). Peraturan Presiden Nomor 72 Tahun 2021 tentang Percepatan Penurunan *Stunting* menyatakan bahwa *stunting* adalah gangguan pertumbuhan dan perkembangan anak yang disebabkan oleh kekurangan gizi kronis serta infeksi berulang. Hal ini ditandai dengan tinggi atau panjang badan anak yang

berada di bawah standar yang ditetapkan oleh Menteri Kesehatan (Ilmu Pemerintahan Suara Khatulistiwa Akreditasi Jurnal Nasional Sinta *et al.*, 2023). Masalah sosial terkait status gizi di masyarakat mencakup *stunting* (anak dengan tinggi badan pendek berdasarkan usia), *underweight* (anak dengan berat badan kurang berdasarkan usia), dan *wasting* (anak dengan berat badan kurang berdasarkan tinggi badan). Penelitian menggunakan metode *K-Nearest Neighbor* untuk menentukan status gizi dengan melibatkan variabel tinggi dan berat badan, menggunakan perhitungan jarak *Euclidean* yang menunjukkan kinerja optimal. Hal ini terlihat dari hasil uji kinerja sistem yang menggunakan jarak *Euclidean*, dengan tingkat akurasi mencapai 100% (Prasetya *et al.*, 2020).

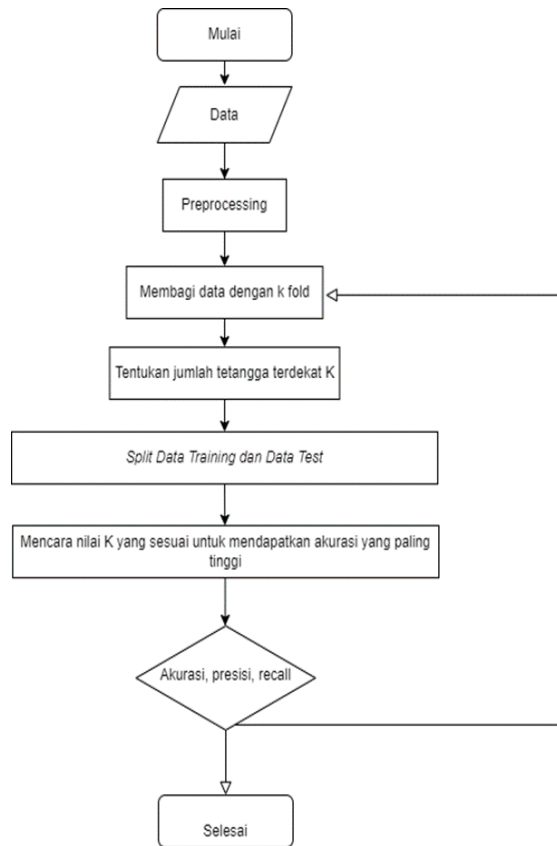
K-Nearest Neighbor (K-NN) adalah algoritma pembelajaran terawasi di mana hasil klasifikasi ditentukan oleh mayoritas tetangga terdekat ke-K. Algoritma ini bekerja dengan menghitung jarak terdekat antara data baru dan tetangga terdekatnya, serta mengandalkan memori dan mayoritas tetangga untuk melakukan klasifikasi. Berdasarkan hal tersebut, penelitian ini akan menggunakan metode Algoritma *K-Nearest Neighbor* untuk mengklasifikasikan status gizi balita berdasarkan kesejahteraan masyarakat (Saeful Bachri *et al.*, 2021). Parameter yang digunakan dalam penelitian ini berdasarkan data antropometrik, yaitu Umur, Berat Badan, dan Tinggi Badan. Pengujian dilakukan dengan perhitungan manual, diikuti dengan pembuatan peringkat, dan kemudian diimplementasikan dalam *Python*. Penelitian sebelumnya yang berkaitan dengan klasifikasi status *stunting* pada balita telah dilakukan oleh beberapa peneliti. Tio Prasetya, Irfan Ali, Cep Lukman Rohmat, dan Odi Nurdiawan (2020) melakukan penelitian berjudul Klasifikasi Status *Stunting* Balita di Desa Slangit Menggunakan Metode *K-Nearest Neighbor*. Penelitian ini menggunakan pendekatan data mining dengan algoritma *K-Nearest Neighbor* (K-NN), di mana perhitungan jarak *Euclidean* digunakan untuk mengklasifikasikan data berdasarkan kelas latih. Parameter yang digunakan adalah data antropometrik, yaitu umur, berat badan, dan tinggi badan. Pengujian dilakukan secara manual dan diimplementasikan ke dalam aplikasi *RapidMiner* untuk mempermudah proses analisis (Prasetya *et al.*,

2020). Selanjutnya, penelitian oleh Otong Saeful Bachri dan Raden Mohamad Herdian Bhakti (2021) berjudul Penentuan Status *Stunting* pada Anak dengan Menggunakan Algoritma KNN juga berfokus pada klasifikasi status gizi balita menggunakan algoritma k-NN. Berdasarkan pengujian terhadap 114 data balita, penelitian ini menunjukkan bahwa algoritma k-NN mampu mengklasifikasikan status *stunting* pada anak dengan memanfaatkan parameter umur, berat badan, tinggi badan, dan lingkaran kepala. Hasil terbaik diperoleh pada nilai $k = 3$, di mana akurasi tertinggi dan error rate terendah dicapai (Saeful Bachri *et al.*, 2021).

Penelitian lainnya dilakukan oleh Ivo Colanus Rally Drajana dan Andi Bode (2022) dalam studi berjudul Prediksi Status Penderita *Stunting* pada Balita Provinsi Gorontalo Menggunakan *K-Nearest Neighbor* Berbasis Seleksi Fitur *Chi Square*. Penelitian ini bertujuan untuk meningkatkan akurasi pendataan *stunting* di Kabupaten Gorontalo, yang sebelumnya dianggap kurang akurat. Dengan menggunakan algoritma k-NN yang digabungkan dengan metode seleksi fitur *Chi Square*, penelitian ini berhasil memprediksi jumlah penderita *stunting* dengan tingkat error rendah, yaitu RMSE 1,200. Penambahan seleksi fitur *Chi Square* terbukti meningkatkan performa algoritma k-NN dibandingkan dengan penggunaan k-NN tanpa seleksi fitur (Colanus *et al.*, 2022). Penelitian-penelitian di atas menunjukkan bahwa algoritma k-NN efektif dalam mengklasifikasikan dan memprediksi status *stunting* balita, terutama ketika dipadukan dengan seleksi fitur yang tepat, seperti *Chi Square*, untuk meningkatkan akurasi dan efisiensi model.

2. Metodologi Penelitian

Metode penelitian yang digunakan mengikuti prosedur standar dengan menerapkan metode *K-Nearest Neighbor*. Pada tahap ini, peneliti akan mengklasifikasikan *stunting* pada balita, dan dengan adanya model ini, peneliti dapat dengan mudah mengaplikasikan algoritma untuk menghasilkan nilai akhir berupa akurasi, presisi, dan *recall*.



Gambar 1. Tahapan Penelitian

Data yang digunakan dalam penelitian ini berasal dari Dinas Kesehatan Kota Salatiga, yaitu data *stunting* balita yang diperoleh pada tahun 2024. Dalam penelitian ini, sangat penting untuk mengidentifikasi variabel yang relevan untuk analisis, seperti variabel *input* (fitur) dan variabel *output* (label) jika menggunakan metode klasifikasi. *Preprocessing* merupakan langkah krusial yang melibatkan berbagai aktivitas untuk membersihkan dan mempersiapkan data agar siap digunakan dalam algoritma. Beberapa langkah umum dalam preprocessing data antara lain:

- 1) **Penanganan Missing Values**
Mengatasi nilai-nilai yang hilang dalam dataset dengan menghapus atau mengisi dengan nilai yang sesuai.
- 2) **Normalisasi**
Mengubah skala data agar memiliki rentang yang sama, biasanya antara 0 dan 1, untuk menghindari fitur dengan rentang besar mendominasi hasil.
- 3) **Encoding**
Mengubah data kategori menjadi numerik menggunakan teknik seperti *one-hot encoding* atau *label encoding*.

4) Pembersihan Data

Menghapus data duplikat atau data yang tidak relevan.

5) *Feature Scaling*

Mengatur skala fitur menggunakan teknik seperti standardisasi untuk memastikan semua fitur berkontribusi setara dalam proses analisis.

Metode *k-fold cross-validation* adalah salah satu cara yang paling efektif untuk menguji performa model dengan menghindari masalah *overfitting*. Pada teknik ini, dataset dibagi menjadi *k* bagian atau lipatan yang berukuran sama. Setiap lipatan digunakan sebagai data uji sekali, sementara sisa lipatan digunakan sebagai data latih. Proses ini diulang sebanyak *k* kali, dan rata-rata dari hasil evaluasi model menjadi indikasi performa model yang lebih akurat. Langkah selanjutnya adalah memilih nilai awal *K* yang akan digunakan dalam algoritma *KNN*.

Pemilihan nilai *K* adalah aspek kritis yang menentukan keakuratan model. Nilai *K* yang terlalu kecil (misalnya, *K* = 1) bisa membuat model terlalu sensitif terhadap data *noise*, sementara nilai *K* yang terlalu besar bisa membuat model menjadi terlalu umum dan kehilangan detail. Pada tahap ini, data dibagi menjadi dua bagian utama: data latih (*training data*) dan data uji (*test data*). Data latih digunakan untuk membangun model, sedangkan data uji digunakan untuk mengukur akurasi prediksi dari model yang telah dibangun. Pembagian ini memungkinkan evaluasi kemampuan model dalam melakukan generalisasi. Setelah membagi data dan memilih beberapa nilai *K* yang berbeda, langkah berikutnya adalah mencari nilai *K* yang memberikan hasil terbaik.

Hasil terbaik biasanya diukur berdasarkan beberapa metrik evaluasi, seperti akurasi, presisi, dan *recall*. Proses ini sering kali melibatkan visualisasi kinerja model pada berbagai nilai *K* untuk memilih parameter yang memberikan keseimbangan antara bias dan varians. Pada tahap evaluasi ini, kita menggunakan beberapa metrik untuk mengukur performa model:

1) Akurasi

Mengukur persentase prediksi yang benar dibandingkan dengan keseluruhan prediksi.

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Prediksi}}$$

2) Presisi

Mengukur kemampuan model untuk menghindari kesalahan positif. Presisi yang tinggi berarti jumlah false positive rendah.

$$\text{Presisi} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

3) Recall

Mengukur kemampuan model untuk menemukan semua contoh positif yang sebenarnya. Recall yang tinggi berarti jumlah false negative rendah.

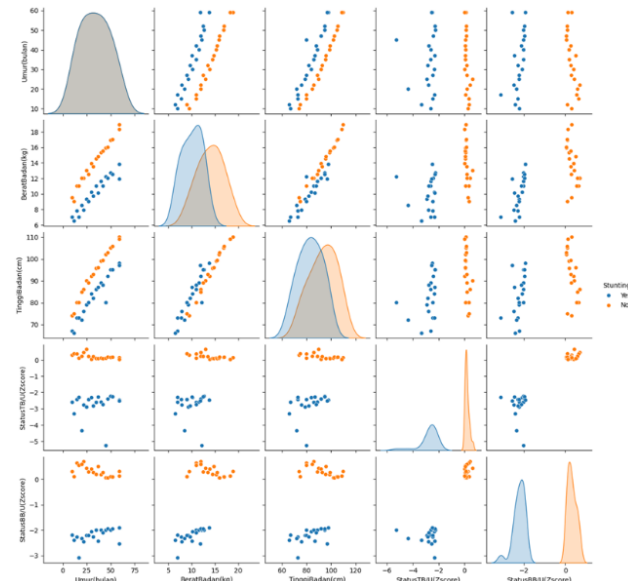
$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Pemilihan nilai K dalam algoritma KNN sangat penting karena secara langsung mempengaruhi kemampuan model dalam mengklasifikasikan data baru. Metode *cross-validation* adalah alat yang efektif untuk memilih nilai K yang optimal, karena memungkinkan evaluasi model pada data yang berbeda-beda dan mengurangi risiko *overfitting*. Tahap *preprocessing* yang baik, termasuk penanganan data dan normalisasi, juga sangat berpengaruh pada hasil akhir. Oleh karena itu, setiap langkah dalam diagram alir memainkan peran krusial dalam membangun model KNN yang andal dan akurat. Dengan menggunakan pendekatan yang sistematis dan berbasis data seperti yang digambarkan dalam diagram, kita dapat mengembangkan model KNN yang mampu membuat prediksi yang lebih baik dan lebih stabil dalam menghadapi data yang beragam.

3. Hasil dan Pembahasan

Hasil

Penelitian ini menggunakan data dari Dinas Kesehatan Kota Salatiga. Data yang digunakan berasal dari hasil penarikan data *stunting* balita pada tahun 2024. Data tersebut memiliki atribut seperti Umur, Jenis Kelamin, Tinggi Badan (TB), Berat Badan (BB), Status BB/U, dan Status TB/U sebagai label. Dalam penelitian ini, digunakan dua data utama: yaitu status TB/U dan status BB/U untuk mengukur apakah balita tersebut mengalami *stunting* atau tidak. Data ini nantinya akan dibagi menjadi data latih (*training*) dan data uji (*testing*).



gambar 2. Grafik Data *Stunting*

Gambar ini merupakan visualisasi *pair plot*, yang menampilkan hubungan antar variabel dalam bentuk *scatter plots* dan distribusi masing-masing variabel dalam bentuk *kde plot* (Kernel Density Estimation). Gambar ini mencakup beberapa variabel yang diukur, yakni Umur (bulan), Berat Badan (kg), Tinggi Badan (cm), Status TB/U (Z-score), dan Status BB/U (Z-score). Selain itu, terdapat kategori tambahan yaitu "*Stunting*", yang dibedakan dengan warna biru dan oranye untuk masing-masing kelompok yang "Yes" (*stunting*) dan "No" (tidak *stunting*). Variabel "Umur" diukur dalam bulan dan terlihat pada bagian pertama grafik *scatter* dalam setiap kolom dan baris. Berdasarkan visualisasi, tampak bahwa Umur berkorelasi positif dengan Berat Badan dan Tinggi Badan.

Hal ini dapat dilihat dari pola *scatter* yang teratur di mana semakin tinggi usia, semakin tinggi pula berat dan tinggi badan individu, baik pada kelompok *stunting* maupun tidak. Namun, pola tersebut terlihat lebih renggang pada kelompok yang mengalami *stunting* (warna biru), yang menunjukkan adanya hubungan yang lebih lemah antara pertambahan usia dan pertumbuhan fisik. Pada kolom kedua, variabel Berat Badan (kg) menunjukkan hubungan kuat dengan Tinggi Badan (cm), terutama pada anak-anak yang tidak mengalami *stunting*. Hal ini terlihat dari grafik yang menunjukkan garis lurus dan teratur, di mana anak-anak dengan berat badan yang lebih tinggi cenderung memiliki tinggi badan yang lebih besar.

Sebaliknya, pada kelompok *stunting*, hubungan ini sedikit lebih longgar, menunjukkan bahwa anak-anak dengan *stunting* mungkin tidak mengalami peningkatan tinggi badan yang sesuai dengan peningkatan berat badan. Pada variabel Tinggi Badan (cm), yang merupakan kolom ketiga, kita dapat melihat pola yang serupa dengan Berat Badan. Ada korelasi yang kuat antara tinggi badan dengan variabel lain seperti Umur dan Berat Badan, terutama pada kelompok non-*stunting*. Selain itu, dari distribusi Tinggi Badan pada plot *kde*, kelompok anak-anak dengan *stunting* memiliki distribusi yang lebih rendah secara umum, menunjukkan bahwa anak-anak yang mengalami *stunting* cenderung memiliki tinggi badan yang lebih rendah daripada yang tidak mengalami *stunting*.

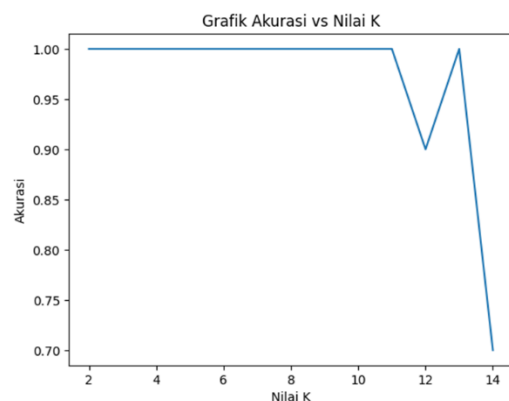
Variabel terakhir yang ditampilkan adalah Status TB/U (Z-Score) dan Status BB/U (Z-Score), yang memberikan indikasi tentang status gizi anak berdasarkan tinggi badan terhadap umur dan berat badan terhadap umur. Anak-anak yang mengalami *stunting* (warna biru) menunjukkan nilai Z-Score yang lebih rendah pada kedua variabel ini, yang berarti mereka berada dalam kategori kurang gizi dibandingkan dengan anak-anak lain yang tidak mengalami *stunting* (warna oranye). Sebaliknya, anak-anak yang tidak mengalami *stunting* memiliki Z-Score yang lebih bervariasi namun lebih tinggi secara rata-rata, menunjukkan status gizi yang lebih baik. Visualisasi ini memberikan wawasan mendalam mengenai hubungan antara *stunting* dan berbagai faktor pertumbuhan fisik seperti berat dan tinggi badan. Dari visualisasi, dapat disimpulkan bahwa anak-anak yang mengalami *stunting* menunjukkan perkembangan fisik yang lebih terbatas dibandingkan dengan yang tidak mengalami *stunting*, terutama terlihat pada distribusi tinggi badan dan Z-Score yang lebih rendah.

	precision	recall	f1-score	support
No	1.00	1.00	1.00	5
Yes	1.00	1.00	1.00	5
accuracy			1.00	10
macro avg	1.00	1.00	1.00	10
weighted avg	1.00	1.00	1.00	10

Gambar 3. Hasil Evaluasi

Gambar 3 ini menampilkan informasi yang menyajikan hasil evaluasi model klasifikasi menggunakan metrik *precision*, *recall*, dan *f1-score*. Terdapat dua kelas dalam klasifikasi ini, yaitu kelas "No" dan kelas "Yes", yang masing-masing mendapatkan dukungan atau *support* sebesar 5 data. Hasil ini juga mencakup *accuracy*, *macro average*, dan *weighted average* dari keseluruhan metrik yang digunakan. Pada kelas "No" dan "Yes", model menunjukkan kinerja yang sempurna dengan nilai *precision*, *recall*, dan *f1-score* semuanya mencapai 1.00. Hal ini berarti bahwa model ini tidak membuat kesalahan dalam mengklasifikasikan kedua kelas tersebut. Nilai *precision* 1.00 menunjukkan bahwa semua prediksi positif yang dilakukan model untuk kedua kelas ini semuanya benar, tanpa ada kesalahan positif (*false positive*). Sementara itu, *recall* 1.00 menunjukkan bahwa model berhasil mengidentifikasi seluruh contoh positif dengan tepat, tanpa ada kesalahan negatif (*false negative*).

Dari keseluruhan metrik, nilai *accuracy* model mencapai 1.00 atau 100%. Nilai *macro average* dan *weighted average* juga sebesar 1.00 untuk masing-masing metrik. *Macro average* menghitung rata-rata sederhana dari *precision*, *recall*, dan *f1-score* dari kedua kelas, tanpa memperhatikan jumlah sampel di tiap kelas. Sementara itu, *weighted average* memperhitungkan kontribusi jumlah sampel di tiap kelas saat menghitung rata-rata, sehingga memberikan bobot yang lebih tinggi kepada kelas dengan jumlah sampel lebih banyak, meskipun dalam hal ini kedua kelas memiliki jumlah sampel yang sama. Nilai-nilai ini menunjukkan bahwa model memiliki performa yang sangat baik dan tidak ada ketidakseimbangan dalam klasifikasi kelas-kelasnya.



Gambar 4. Grafik Akurasi

Gambar di atas menunjukkan grafik yang mengilustrasikan hubungan antara nilai K dengan *accuracy* pada algoritma *K-Nearest Neighbors* (KNN). Sumbu X mewakili nilai K yang digunakan dalam model KNN, dan sumbu Y mewakili tingkat *accuracy* yang diperoleh. Pada grafik ini, terlihat adanya fluktuasi *accuracy* saat nilai K berubah. Pada awal grafik, untuk nilai K yang lebih kecil (2 hingga sekitar 8), *accuracy* berada pada titik maksimum di 100%. Ini menunjukkan bahwa untuk nilai-nilai K yang kecil, model KNN memberikan kinerja yang sangat baik pada data yang digunakan. Namun, setelah K mencapai nilai 10, *accuracy* mulai menurun tajam, menandakan bahwa model mungkin menjadi terlalu sederhana atau tidak cukup robust ketika K meningkat lebih besar dari 8 atau 10. Di sekitar nilai K antara 10 dan 14, kita melihat fluktuasi *accuracy* yang signifikan, di mana *accuracy* naik dan turun dengan cepat. Penurunan yang sangat drastis terjadi pada $K = 14$, yang menunjukkan bahwa model KNN pada nilai K tersebut menghasilkan prediksi yang kurang akurat, mungkin karena K terlalu besar, sehingga model gagal menangkap pola penting dari data.

Pembahasan

Dalam penelitian ini, penerapan algoritma *K-Nearest Neighbors* (KNN) untuk mendeteksi status *stunting* pada balita menunjukkan hasil yang signifikan, sejalan dengan temuan dari penelitian sebelumnya seperti yang dilakukan oleh Colanus *et al.* (2022) dan Prasetya *et al.* (2020). Visualisasi *pair plot* yang digunakan untuk menganalisis hubungan antar variabel seperti usia, berat badan, dan tinggi badan menunjukkan adanya korelasi positif antara usia dengan berat badan dan tinggi badan pada balita yang tidak mengalami *stunting*. Sebaliknya, pada kelompok balita yang mengalami *stunting*, korelasi ini cenderung lebih lemah, mengindikasikan pertumbuhan fisik yang terbatas, yang sesuai dengan penelitian Rahmawati *et al.* (2019) yang menemukan bahwa pengetahuan orang tua mengenai *stunting* sangat mempengaruhi pencegahan gangguan pertumbuhan pada anak. Preprocessing data, seperti penanganan *missing values*, normalisasi data, dan *feature scaling*, terbukti meningkatkan kinerja model KNN, serupa dengan hasil yang diperoleh oleh Saeful Bachri *et al.* (2021), yang menunjukkan bahwa data yang bersih dan terstandarisasi meningkatkan akurasi model dalam mengklasifikasikan status *stunting*.

Hasil evaluasi model menggunakan metrik *precision*, *recall*, dan *f1-score* menunjukkan kinerja yang sangat baik, dengan nilai 1.00 untuk kedua kelas "*stunting*" dan "*tidak stunting*". Ini menandakan bahwa model berhasil mengidentifikasi balita yang mengalami *stunting* dan tidak dengan sangat akurat, tanpa kesalahan prediksi (*false positive* atau *false negative*). Hasil ini juga mendukung temuan Colanus *et al.* (2022), yang menggunakan seleksi fitur untuk meningkatkan akurasi dalam klasifikasi status *stunting* dengan menggunakan algoritma KNN. Namun, pemilihan nilai K yang optimal menjadi tantangan utama, karena pada nilai K kecil (2 hingga 8), model memberikan akurasi yang sangat baik, tetapi ketika K meningkat lebih dari 10, performa mulai menurun tajam. Fenomena ini mengindikasikan bahwa model menjadi terlalu sederhana atau terlalu kompleks, yang menyebabkan masalah seperti *overfitting* atau *underfitting*, sebagaimana dijelaskan oleh Ridho Al Izzati & Elza Elmira (2019). Oleh karena itu, pemilihan nilai K yang tepat sangat penting dalam menjaga keseimbangan antara ketepatan prediksi dan kompleksitas model.

Penelitian ini menunjukkan bahwa algoritma KNN dapat digunakan secara efektif untuk mendeteksi status *stunting* pada balita dengan data yang relatif sederhana dan preprocessing yang tepat. Hasil yang diperoleh memberikan akurasi yang sangat baik, namun penting untuk mempertimbangkan teknik tambahan seperti *cross-validation* untuk meningkatkan ketahanan model terhadap variasi data dan memastikan generalisasi yang lebih baik. Selain itu, faktor sosial-ekonomi, seperti yang dibahas oleh Mentari & Hermansyah (n.d.), turut berperan dalam status gizi balita. Oleh karena itu, memasukkan faktor-faktor ini dalam model prediksi dapat meningkatkan akurasi model lebih lanjut dan memberikan hasil yang lebih komprehensif.

4. Kesimpulan

Kesimpulan dari penelitian ini menunjukkan bahwa algoritma *K-Nearest Neighbors* (KNN) berhasil menghasilkan tingkat akurasi, presisi, dan *recall* yang sangat tinggi, dengan nilai sebesar 100%. Hal ini membuktikan bahwa model klasifikasi yang digunakan mampu mengidentifikasi seluruh balita

yang mengalami *stunting* dengan tepat, tanpa kesalahan dalam klasifikasi. Grafik yang menggambarkan hubungan antara nilai K dan akurasi menunjukkan bahwa nilai K yang kecil (sekitar 2 hingga 8) memberikan akurasi yang tinggi. Namun, seiring dengan meningkatnya nilai K , akurasi mulai menurun, yang menandakan bahwa nilai K yang terlalu besar dapat mengurangi keakuratan model dalam mengklasifikasikan status *stunting*. Penelitian ini memberikan kontribusi penting dalam bidang klasifikasi data mining, khususnya dalam deteksi *stunting* pada balita. Model ini berpotensi besar untuk membantu tenaga kesehatan dalam mendeteksi *stunting* secara lebih akurat dan cepat, yang sangat penting dalam melakukan intervensi yang tepat sasaran. Deteksi dini *stunting* pada balita sangat penting untuk memastikan intervensi yang lebih cepat dan efektif, dan algoritma KNN terbukti menjanjikan dalam mendukung proses tersebut. Pemanfaatan teknologi seperti *machine learning* dapat memainkan peran krusial dalam upaya pencegahan *stunting* dengan cara yang lebih efisien dan akurat, mendukung tujuan kesehatan masyarakat untuk meningkatkan status gizi anak-anak di Indonesia.

5. Daftar Pustaka

- Bachri, O. S., & Bhakti, R. M. H. (2021). Penentuan Status Stunting pada Anak dengan Menggunakan Algoritma KNN. *Jurnal Ilmiah Intech: Information Technology Journal of UMUS*, 3(02), 130-137. <https://doi.org/10.46772/intech.v3i02.533>.
- Drajana, I. C. R., & Bode, A. (2022). Prediksi Status Penderita Stunting Pada Balita Provinsi Gorontalo Menggunakan K-Nearest Neighbor Berbasis Seleksi Fitur Chi Square. *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, 5(2).
- Fitrauni, R., Muchlis, N., & Arman, A. (2022). Strategi percepatan penurunan stunting di Kabupaten Sigi. *Journal of Muslim Community Health*, 3(4), 193-209. <https://doi.org/10.52103/jmch.v3i4.1237>.
- Husaini, A. L., Hoeronis, I., Lumana, H. H., & Puspahreni, L. D. (2023). Early Detection of Stunting in Toddlers Based on Ensemble Machine Learning in Purbaratu Tasikmalaya. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 11(3), 487-495.
- Izwardy, D. (2019). Kebijakan dan strategi penanggulangan stunting di Indonesia. *Kementerian Kesehatan Republik Indonesia*.
- Kamila, R., Novita, C., Harahap, Y. A., & Ula, M. (2024, November). DETEKSI STUNTING PADA BALITA MENGGUNAKAN METODE K-NEAREST NEIGHBOR. In *Prosiding Seminar Nasional Teknologi dan Teknik Informatika (SENASTIKA)* (Vol. 1, No. 1).
- Kemenkes, R. I. (2023). *Prevalensi Stunting di Indonesia Turun ke 21, 6% dari 24, 4%*. Kemenkes.
- Lonang, S., Yudhana, A., & Biddinika, M. K. (2023). Analisis Komparatif Kinerja Algoritma Machine Learning untuk Deteksi Stunting. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(4), 2109-2117.
- Mentari, S., & Hermansyah, A. (2019). Faktor-faktor yang berhubungan dengan status stunting anak usia 24-59 bulan di wilayah kerja UPK puskesmas Siantan Hulu. *Pontianak Nutrition Journal (PNJ)*, 1(1), 1-5. <https://doi.org/10.30602/pnj.v1i1.275>.
- Prasetya, T., Ali, I., Rohmat, C. L., & Nurdyawan, O. (2020). Klasifikasi Status Stunting Balita Di Desa Slangit Menggunakan Metode K-Nearest Neighbor. *INFORMATICS FOR EDUCATORS AND PROFESSIONAL: Journal of Informatics*, 5(1), 93-104.
- Priyono, P. (2020). Strategi Percepatan Penurunan Stunting Perdesaan (Studi Kasus Pendampingan Aksi Cegah Stunting di Desa Banyumundu, Kabupaten Pandeglang). *Jurnal Good Governance*. <https://doi.org/10.32834/gg.v16i2.198>.

- Rahman, H., Rahmah, M., & Saribulan, N. (2023). Upaya Penanganan Stunting di Indonesia: Analisis Bibliometrik Dan Analisis Konten. *Jurnal Ilmu Pemerintahan Suara Khatulistiwa*, 8(1), 44-59. <https://doi.org/10.33701/jipsk.v8i1.3184>.
- Rahmawati, A., Nurmawati, T., & Permata Sari, L. (2019). Faktor yang berhubungan dengan pengetahuan orang tua tentang stunting pada balita. *Jurnal Ners Dan Kebidanan (Journal of Ners and Midwifery)*, 6(3), 389–395. <https://doi.org/10.26699/jnk.v6i3.art.p389-395>.
- Sany, F. A. (2023). Penerapan Sistem Pakar untuk Deteksi Stunting. *Jurnal Ilmiah Ecosystem*, 23(3), 602-609. <https://doi.org/10.35965/eeco.v23i3.3774>.
- Satriawan, E. (2018). Strategi nasional percepatan pencegahan stunting 2018-2024. *Jakarta: Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K)*.