Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.5, October 2025, pp. 2711~2717
e-ISSN: 2548-6861

Comparison of KNN and Naïve Bayes Classification Algorithms for Predicting Stunting in Toddlers in Banjaran District

Rivaldy Fauzan Mu'taz [1]*, Ai Rosita [2]*
* Teknik Informatika, Universitas Widyatama
rivaldy.fauzan@widyatama.ac.id [1], ai.rosita@widyatama.ac.id [2]

Article Info

Article history:

Received 2025-06-02
Revised 2025-08-23
Accepted 2025-09-03

Keyword:

Stunting,
Machine Learning,
Classification,
Naïve Bayes,
K-Nearest Neighbors (KNN),
Data Pre-processing,
SMOTE.

ABSTRACT

Stunting is a chronic nutritional problem that seriously impacts child growth and development. This study aims to compare the performance of the Naïve Bayes and K-Nearest Neighbors (KNN) algorithms in predicting stunting in toddlers in Banjaran District. The dataset consists of 12,000 toddler data points with three main features: age, gender, and height. The research employed a quantitative approach by applying machine learning algorithms. The SMOTE oversampling technique was applied only to the training data to avoid data leakage, and 5-fold cross-validation was used. A K-value of 3 was selected for the final KNN model based on validation curve analysis to prevent overfitting. The results show that KNN significantly outperformed Naïve Bayes across all evaluation metrics. The Naïve Bayes model yielded an accuracy of 67.50%, precision of 50.87%, recall of 61.38%, F1-score of 55.63%, specificity of 70.54%, and an AUC score of 75.71%. Meanwhile, the KNN (K=3) model achieved an accuracy of 99.11%, precision of 98.08%, recall of 99.25%, F1-score of 98.66%, specificity of 99.03%, and an AUC score of 99.65%. The performance difference between the two models was confirmed by McNemar's Test with a p-value < 0.05, indicating a statistically significant difference. The low performance of Naïve Bayes was attributed to the violation of the feature independence assumption, particularly the high correlation between age and height ($r \approx 0.87$). In conclusion, KNN is the more appropriate algorithm for stunting prediction on this dataset. However, the limitation of features suggests the need for further research with additional variables and external validation before wider-scale implementation.

I. PENDAHULUAN

Stunting, kondisi gagal tumbuh pada balita yang ditandai tinggi badan lebih pendek dari standar usianya, merupakan masalah gizi kronis signifikan. Balita dikategorikan stunting jika standar pertumbuhannya berada di bawah -2 standar deviasi (SD) dari standar pertumbuhan anak WHO [9]. Prevalensi stunting di Indonesia masih tinggi, mencapai 21,5% pada tahun 2023 menurut Survei Status Gizi Indonesia (SSGI) [6], melampaui ambang batas WHO. Kondisi ini menuntut upaya deteksi dini yang lebih efektif untuk intervensi tepat waktu.

Dalam upaya meningkatkan efektivitas deteksi dini tersebut, pendekatan komputasional berbasis data, khususnya menggunakan teknik *machine learning*, menawarkan alternatif yang menjanjikan. *Machine learning* memiliki kemampuan untuk mempelajari pola dari data historis dan membuat prediksi atau klasifikasi terhadap data baru.

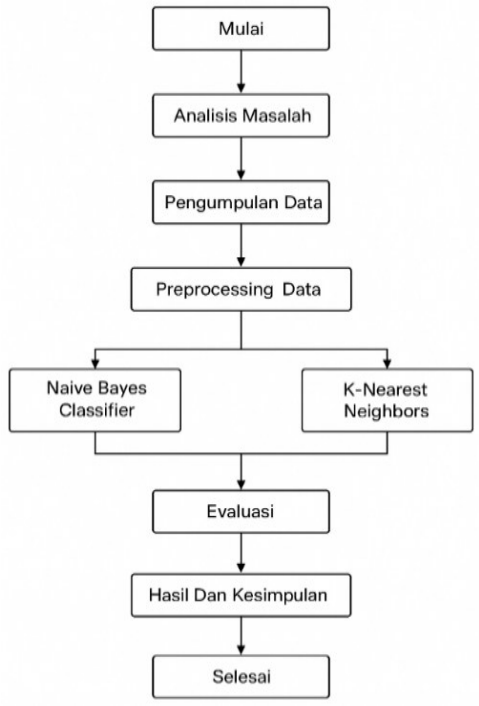
Penelitian ini berfokus pada perbandingan dua algoritma klasifikasi populer, yaitu *Naïve Bayes Classification* dan *K-Nearest Neighbors* (KNN), untuk memprediksi potensi stunting pada balita di Kecamatan Banjaran. Penelitian sebelumnya menunjukkan bahwa kedua algoritma ini telah diterapkan dalam konteks prediksi stunting dengan hasil yang bervariasi. Wahyudina dkk. [22] menunjukkan kapabilitas *Naïve Bayes* dalam menghasilkan akurasi yang baik (85,33%) di Semarang. Drajana & Bode [23] menyoroti penerapan KNN di Gorontalo dan menunjukkan bahwa performanya

dapat ditingkatkan melalui penggunaan seleksi fitur. Sementara itu, studi Fannany & Gunawan [19] di Bojongsoang, meskipun menemukan *Random Forest* lebih unggul, tetap menunjukkan *Naïve Bayes* sebagai metode yang kompetitif dan menggarisbawahi pentingnya teknik penanganan data tidak seimbang seperti *oversampling*.

Tujuan dari penelitian ini adalah (1) membangun model prediksi stunting pada balita menggunakan algoritma klasifikasi *Naïve Bayes* dan *K-Nearest Neighbors* (KNN), dan (2) mengevaluasi serta membandingkan performa (akurasi dan metrik evaluasi lainnya) dari kedua model prediksi stunting tersebut. Penelitian ini diharapkan dapat memberikan manfaat berupa pengembangan pemahaman tentang aplikasi *machine learning* di bidang kesehatan, khususnya deteksi dini stunting; memberikan bukti empiris mengenai perbandingan kinerja algoritma *Naïve Bayes* dan KNN untuk kasus prediksi stunting; memberikan masukan berbasis data kepada pembuat kebijakan dan tenaga kesehatan di Kecamatan Banjaran mengenai potensi penerapan algoritma ini sebagai bahan pertimbangan dalam mengembangkan solusi teknologi deteksi dini stunting; serta memberikan kontribusi bagi pengembangan ilmu pengetahuan di bidang informatika kesehatan dan teknologi *machine learning.*

II. METODE

Pelaksanaan penelitian melibatkan beberapa tahapan utama sebagaimana dapat dilihat pada gambar 1, yaitu: analisis masalah, pengumpulan data, *preprocessing* data, implementasi dan pelatihan model, evaluasi model, hingga penarikan hasil dan kesimpulan.



Gambar 1. Metodologi Penelitian

A. Analisis Masalah

Tahap awal melibatkan pemahaman mendalam mengenai permasalahan stunting, didukung oleh studi literatur untuk mengidentifikasi faktor-faktor risiko dan pendekatan yang telah ada. Penentuan status gizi, khususnya stunting, didasarkan pada standar pertumbuhan WHO [9]. Standar ini menggunakan Z-score, yang merupakan ukuran statistik yang menunjukkan seberapa jauh suatu nilai pengukuran individu (misalnya, tinggi badan anak) dari nilai tengah (median) populasi referensi yang sehat dan sesuai usia, dinyatakan dalam satuan standar deviasi (SD). Z-score dihitung menggunakan rumus berikut:

$$Z\ score = \frac{TB\ Hitung\ -\ \mu}{\sigma}$$

Keterangan dari formula tersebut adalah:

- Z-Score = Deviasi nilai individu dari nilai rata-rata median dibagi dengan standar deviasi referensi
- TB Hitung = Nilai dari hasil pengukuran tinggi badan balita
- μ\mu = Median Nilai Populasi Referensi/ Rata-rata tinggi anak seusia (WHO *Growth Standard*)
- σ\sigma = Standar Deviasi Populasi Referensi / Hasil selisih nilai standar deviasi dari kasus yang dihitung

Interpretasi Z-score Tinggi Badan menurut Umur (TB/U) untuk klasifikasi status gizi mengacu pada standar WHO, seperti dirangkum pada Tabel I.

TABEL I
INTERPRETASI Z-SCORE TB/U BERDASARKAN WHO

Kategori Status Gizi	Rentang Z-Score	Keterangan
Severely Stunting	Z < -3 SD	Anak mengalami stunting sangat parah
Stunting	-3 SD ≤ Z < -2 SD	Anak tergolong stunting/pendek
Normal	-2 SD ≤ Z ≤ +2 SD	Pertumbuhan anak normal/sehat
Tinggi (*Tall*)	Z > +2 SD	Pertumbuhan anak berada di atas rata-rata

Berdasarkan analisis masalah dan definisi stunting tersebut, pendekatan *machine learning* dipilih sebagai metode untuk mengembangkan model prediksi. Algoritma klasifikasi *Naïve Bayes* dan *K-Nearest Neighbors* (KNN) dipilih untuk dibandingkan karena potensi kinerjanya yang telah ditunjukkan dalam penelitian terkait stunting sebelumnya [22][23], serta karena keduanya merepresentasikan pendekatan klasifikasi yang berbeda (probabilistik untuk *Naïve Bayes* dan *instance based* untuk KNN) yang menarik untuk dievaluasi pada *dataset* lokal.

B. Pengumpulan Data

Data primer berupa dataset balita (usia 0-60 bulan) diperoleh dari Seksi Pemberdayaan Masyarakat Kecamatan Banjaran, mencakup sampel 12.000 data dari dua puskesmas. *Dataset* berisi informasi pengukuran antropometri dan status gizi balita yang telah ditentukan berdasarkan standar pertumbuhan WHO dan interpretasi Z-score sebagaimana dijelaskan pada Tabel I. Struktur dataset yang digunakan dalam penelitian ini dirangkum pada Tabel II.

TABEL II
STRUKTUR DATASET

Kolom	Tipe Data	Keterangan
umur_bulan	Integer	Usia balita dalam bulan (0–60 bulan)
jenis_kelamin	String	Jenis kelamin balita (laki-laki/perempuan)
tinggi_badan	Float	Tinggi badan balita dalam centimeter
status_gizi	String	Kategori status gizi: severely stunting, stunting, normal, tinggi

Fitur utama yang digunakan meliputi Umur (bulan), Jenis Kelamin, dan Tinggi Badan (cm). Variabel target 'Status Gizi' asli (Normal, Stunting, *Severely Stunted*, Tinggi) ditransformasi menjadi variabel target biner 'Status Gizi Biner' (0: Normal, 1: Stunting) setelah mengeksklusi 1.940 data dengan kategori 'Tinggi', menghasilkan 10.060 data untuk pemodelan.

C. Pra-pemrosesan Data

Tahap ini krusial untuk menyiapkan data mentah. Proses yang dilakukan meliputi: (1) Pembersihan dan Penanganan Data Hilang: Analisis awal menunjukkan tidak ada data hilang pada kolom fitur relevan dalam sampel yang digunakan. (2) Transformasi Fitur: Fitur 'Jenis Kelamin' diubah menjadi numerik menggunakan *Label Encoding* (['laki-laki', 'perempuan'] -> [0, 1]), dan fitur numerik 'Umur (bulan)' serta 'Tinggi Badan (cm)' diskalakan menggunakan *StandardScaler*. (3) Penanganan Data Tidak Seimbang: Teknik *oversampling* SMOTE *(Synthetic Minority Over-sampling Technique)* [18] diterapkan hanya pada data latih untuk menyeimbangkan distribusi kelas. (4) Pembagian Data: Setelah data dengan kategori 'Tinggi' sebanyak 1.940 baris dieliminasi, diperoleh dataset final sebanyak 10.060 baris. Dataset ini kemudian dibagi secara terstratifikasi menjadi 80% data latih (8.048 sampel) dan 20% data uji (2.012 sampel). Setelah dilakukan teknik SMOTE pada data latih, jumlah sampel meningkat menjadi 10.752 dengan distribusi kelas yang seimbang.

D. Implementasi dan Pelatihan Model

Dua algoritma klasifikasi diterapkan menggunakan bahasa pemrograman *Python* dan *library Scikit-learn* serta *Imbalanced-learn*: (1) *Naïve Bayes Classifier*: Menggunakan implementasi *Gaussian Naïve Bayes*. (2) *K-Nearest Neighbors (KNN)*: Menggunakan metrik jarak *Euclidean*. Nilai K optimal ditentukan melalui *GridSearchCV* dengan 5-*fold cross-validation* pada data latih (dengan SMOTE) dengan menguji K ganjil dari 1 hingga 21. Berdasarkan analisis risiko *overfitting* pada kurva validasi, nilai K yang lebih *robust* (K=3) dipilih untuk model final, meskipun *GridSearchCV* menunjukkan K=1 sebagai yang tertinggi secara numerik.

E. Evaluasi Model

Kinerja model dievaluasi menggunakan data uji untuk mengukur kemampuan generalisasi. Metrik evaluasi yang digunakan mencakup *Confusion Matrix* untuk mengidentifikasi jenis kesalahan prediksi (TN, FP, FN, TP), serta akurasi untuk mengukur proporsi keseluruhan prediksi yang benar. Evaluasi juga difokuskan pada kelas stunting melalui metrik presisi, recall, dan F1-*score*, guna menilai kemampuan model dalam mengenali kasus positif secara tepat dan menyeluruh. Selain itu, spesifisitas digunakan untuk mengukur kemampuan model dalam mengenali kasus normal dengan benar, sedangkan *Area Under the ROC Curve* (AUC) digunakan untuk menilai kemampuan diskriminatif model di berbagai ambang batas keputusan. Untuk memastikan signifikansi perbedaan performa antar model, dilakukan pula Uji *McNemar* sebagai validasi statistik.

F. Hasil dan Kesimpulan

Tahap akhir dari metodologi ini adalah melakukan analisis komparatif terhadap hasil evaluasi kinerja kedua model (*Naïve Bayes* dan KNN) yang diperoleh pada tahap sebelumnya. Analisis difokuskan pada identifikasi algoritma yang menunjukkan performa prediksi stunting terbaik berdasarkan metrik-metrik yang telah ditetapkan, terutama *Recall* dan F1-*Score* untuk kelas stunting. Hasil analisis ini kemudian menjadi dasar untuk menarik kesimpulan penelitian terkait perbandingan efektivitas kedua algoritma dalam konteks *dataset* yang digunakan dan merumuskan saran lebih lanjut.

III. HASIL DAN PEMBAHASAN

Disajikan hasil kuantitatif dari pra-pemrosesan, analisis diagnostik dan evaluasi kinerja model *Naïve Bayes* dan KNN pada 2.012 data uji (33.20% Stunting, 66.80% Normal).

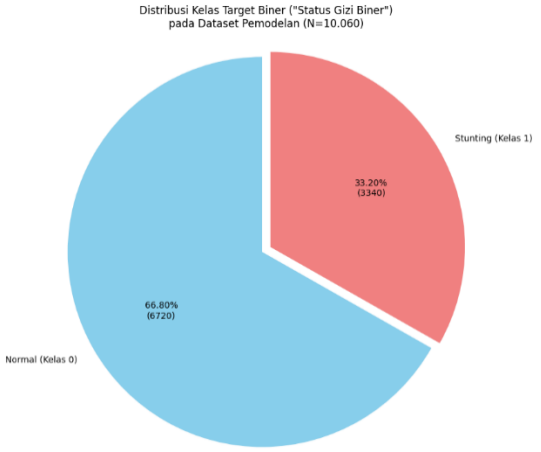
A. Hasil Pra-pemrosesan Data

Distribusi awal kategori 'Status Gizi' pada 12.000 data sampel disajikan pada Tabel III.

TABEL III
DISTRIBUSI AWAL KATEGORI 'STATUS GIZI' PADA SAMPEL DATA
(N=12.000)

Status Gizi	Jumlah	Persentase (%)
Normal	6720	56.00
Severely Stunted	1970	16.42
Tinggi	1940	16.17
Stunted	1370	11.42
Total	**12000**	**100.01**

Setelah eksklusi kategori 'Tinggi' dan transformasi menjadi target biner, distribusi kelas pada 10.060 data adalah 66.80% Normal (Kelas 0) dan 33.20% Stunting (Kelas 1). Visualisasi distribusi ini dapat dilihat pada Gambar 2.



Gambar 2. Distribusi Kelas Target Biner ("Status Gizi Biner") pada Dataset Pemodelan (N=10.060)

Penskalaan fitur numerik menggunakan StandardScaler pada data latih (n=8.048) menghasilkan data dengan rata-rata mendekati 0 dan standar deviasi mendekati 1, seperti dirangkum pada Tabel IV.

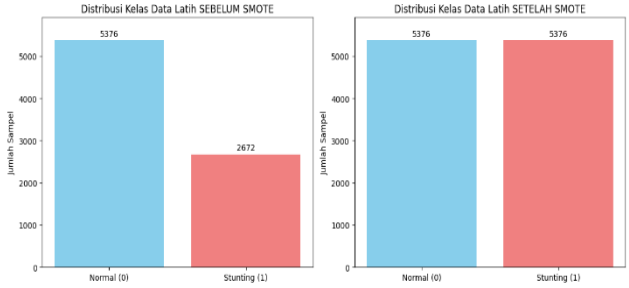
TABEL IV
STATISTIK DESKRIPTIF FITUR NUMERIK SEBELUM DAN SESUDAH STANDARDSCALER

Fitur	Status Skaling	Mean	Std Dev	Min	Max
Umur (bulan)	Sebelum	31.37	17.29	0.00	60.00
	Sesudah	0.00	1.00	-1.82	1.66
Tinggi Badan (cm)	Sebelum	87.43	16.88	40.10	123.60
	Sesudah	0.00	1.00	-2.80	2.14

Data latih yang semula tidak seimbang kemudian diseimbangkan menggunakan teknik SMOTE (Synthetic Minority Over-sampling Technique). Setelah proses ini, jumlah data latih bertambah menjadi 10.752 sampel, dengan distribusi kelas yang merata: 50% untuk kelas Normal dan 50% untuk kelas Stunting. Rincian distribusi sebelum dan sesudah SMOTE dapat dilihat pada Tabel V, sementara visualisasi perbandingannya ditampilkan pada Gambar 3.

TABEL V
DISTRIBUSI KELAS PADA DATA LATIH SEBELUM DAN SESUDAH SMOTE

Kelas	Jumlah Sebelum	Persen Sebelum (%)	Jumlah Sesudah	Persen Sesudah (%)
Normal (0)	5.376	66.8	5.376	50
Stunting (1)	2.672	33.2	5.376	50
Total	**8.048**	**100**	**10.752**	**100**



Gambar 3. Distribusi Kelas pada Data Latih Sebelum dan Sesudah SMOTE

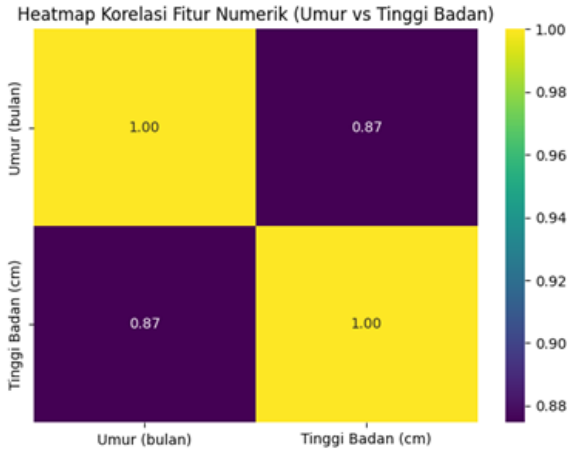
Untuk mengukur pengaruh teknik *oversampling* ini, dilakukan perbandingan kinerja model yang dilatih dengan dan tanpa SMOTE. Hasil evaluasi kinerja masing-masing model dirangkum pada Tabel VI berikut.

TABEL VI
PERBANDINGAN KINERJA MODEL DENGAN DAN TANPA SMOTE

Skenario	Akurasi (%)	Presisi (Stunting) (%)	Recall (Stunting) (%)	F1-Score (Stunting) (%)
NB (Tanpa SMOTE)	69,09	54,26	43,86	48,51
NB (Dengan SMOTE)	67,50	50,87	61,38	55,63
KNN K=3 (Tanpa SMOTE)	99,25	98,80	98,95	98,88
KNN K=3 (Dengan SMOTE)	99,11	98,08	99,25	98,66

B. Analisis Model Diagnostik dan Validasi

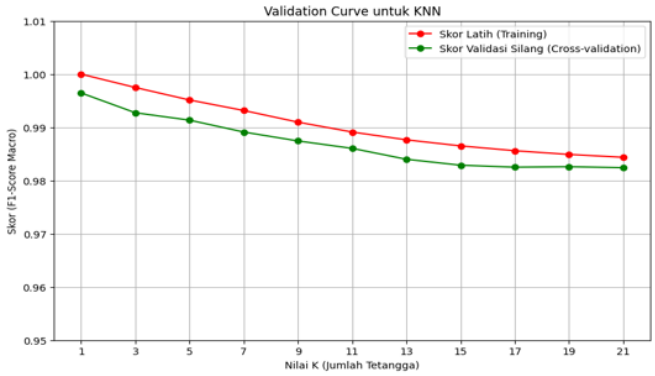
Analisis korelasi Pearson dilakukan untuk menguji asumsi independensi pada Naïve Bayes. Hasilnya disajikan pada Gambar 4.



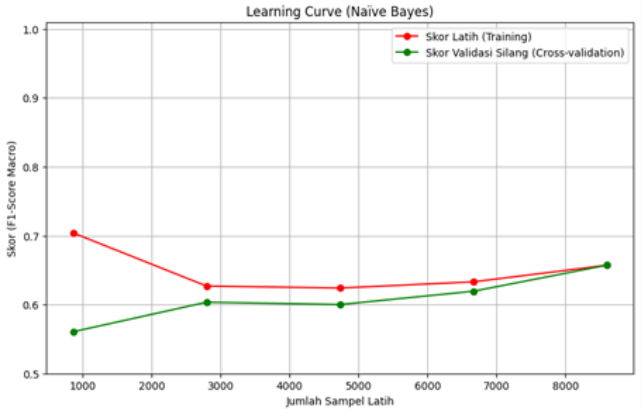
Gambar 4. Heatmap Korelasi Fitur Numerik

Heatmap pada Gambar 4 menunjukkan adanya korelasi positif yang sangat kuat (r≈0.87) antara variabel 'Umur (bulan)' dan 'Tinggi Badan (cm)'. Nilai korelasi yang tinggi ini menandakan bahwa kedua variabel tersebut saling terkait erat; secara logis, seiring bertambahnya umur seorang anak, tinggi badannya pun akan cenderung meningkat. Keterkaitan ini mengindikasikan pelanggaran signifikan terhadap asumsi independensi yang menjadi dasar dari algoritma *Naïve Bayes*. Dengan adanya dependensi yang kuat antara umur dan tinggi badan, pelanggaran asumsi ini berpotensi menurunkan akurasi dan keandalan hasil klasifikasi yang diberikan oleh model.

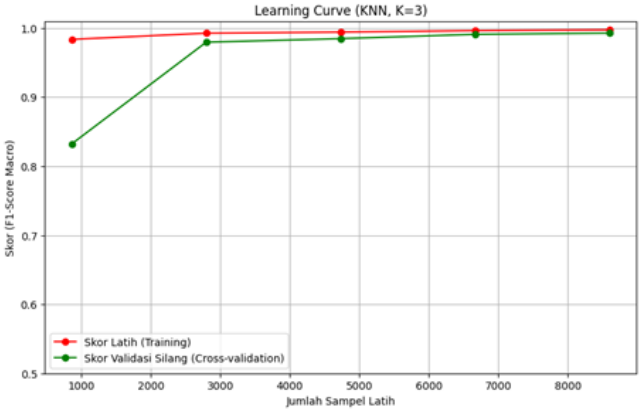
Kemudian Analisis dilanjutkan dengan Kurva Validasi dan Pembelajaran. Untuk menganalisis perilaku model dan risiko *overfitting*, dibuat kurva validasi untuk KNN serta kurva pembelajaran untuk kedua model. Kurva validasi untuk KNN (Gambar 5) menunjukkan bahwa meskipun K=1 memberikan skor validasi silang tertinggi,skor latih (merah) berada di titik sempurna, yaitu 1.0 mengindikasikan risiko *overfitting*. Nilai K=3 dipilih karena menunjukkan keseimbangan yang lebih baik antara kinerja tinggi dan generalisasi. Kurva pembelajaran untuk *Naïve Bayes* (Gambar 6) menunjukkan pola *high bias*, sementara kurva untuk KNN K=3 (Gambar 7) menunjukkan pola model yang *well-fitted*.



Gambar 5. *Validation Curve* untuk KNN



Gambar 6. *Learning Curve* untuk *Naïve Bayes*



Gambar 7. *Learning Curve* untuk KNN (K=3)

C. Kinerja Model Naïve Bayes

Model *Gaussian Naïve Bayes* yang telah dilatih pada data latih yang telah diseimbangkan, dievaluasi menggunakan 2.012 data uji. Hasil *Confusion Matrix* menunjukkan: *True Negative* (TN) = 948, *False Positive* (FP) = 396, *False Negative* (FN) = 258, dan *True Positive* (TP) = 410. Metrik evaluasi yang diperoleh meliputi: akurasi sebesar 67,50%, presisi (Stunting) 50,87%, *recall* (Stunting) 61,38%, F1-*score* (Stunting) 55,63%, spesifisitas (Normal) 70,54%, dan AUC score 75,71%. Hasil lengkap disajikan pada Tabel VII.

TABEL VII
METRIK EVALUASI KINERJA MODEL *NAÏVE BAYES*

Metrik Evaluasi	Nilai (%)
Akurasi	67.50
Presisi (Kelas Stunting)	50.87
Recall (Kelas Stunting)	61.38
F1-Score (Kelas Stunting)	55.63
Spesifisitas (Kelas Normal)	70.54
AUC Score	75.71

D. Kinerja Model K-Nearest Neighbors (KNN)

Nilai K=3 dipilih sebagai model final yang lebih robust. Model KNN dengan K = 3 kemudian dilatih dan dievaluasi menggunakan data uji.

Hasil *Confusion Matrix* menunjukkan *True Negative* (TN) = 1.331, *False Positive* (FP) = 13, *False Negative* (FN) = 5, dan *True Positive* (TP) = 663. Metrik evaluasi yang dihasilkan meliputi: akurasi sebesar 99,11%, presisi (Stunting) 98,08%, *recall* (Stunting) 99,25%, F1-*score* (Stunting) 98,66%, spesifisitas (Normal) 99,03%, dan AUC score 99,65%. Hasil lengkap disajikan pada Tabel VIII.

TABEL VIII
METRIK EVALUASI KINERJA MODEL KNN

Metrik Evaluasi	Nilai (%)
Akurasi	99.11
Presisi (Kelas Stunting)	98.08
Recall (Kelas Stunting)	99.25
F1-*Score* (Kelas Stunting)	98.66
Spesifisitas (Kelas Normal)	99.03
AUC *Score*	99.65

E. Perbandingan Kinerja Model

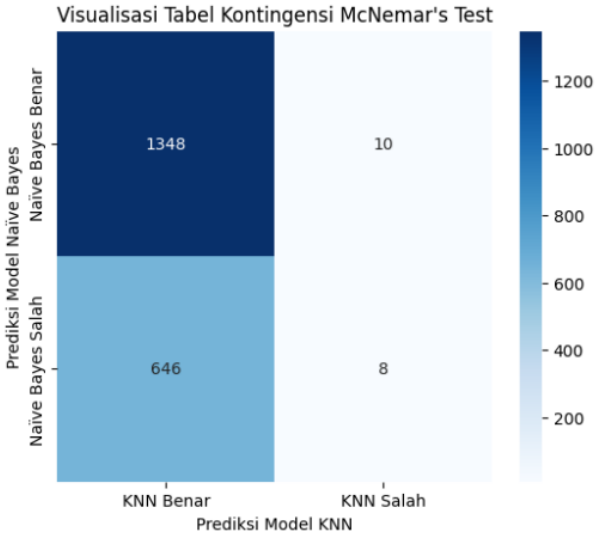
Perbandingan langsung antara kedua model disajikan pada Tabel IX.

TABEL IX
PERBANDINGAN METRIK EVALUASI KINERJA *NAÏVE BAYES* VS KNN

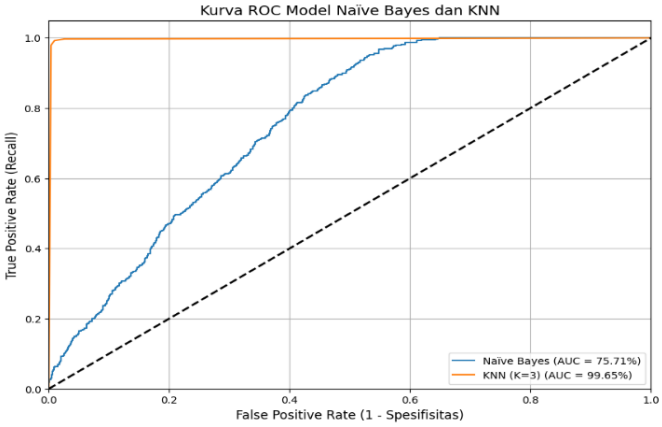
Metrik Evaluasi	*Naïve Bayes* (%)	KNN (K=1) (%)
Akurasi	67.50	99.11
Presisi (Kelas Stunting)	50.87	98.08
Recall (Kelas Stunting)	61.38	99.25
F1-*Score* (Kelas Stunting)	55.63	98.66
Spesifisitas (Kelas Normal)	70.54	99.03
AUC *Score*	75.71	99.65

Berdasarkan Tabel IX, terlihat jelas bahwa model KNN (K=3) secara konsisten mengungguli *Naïve Bayes* di semua metrik evaluasi. Secara khusus, metrik yang paling krusial untuk deteksi dini, yaitu *Recall* (Stunting) dan F1-*Score* (Stunting), menunjukkan perbedaan yang sangat besar, di mana KNN (K=3) mencapai 99.25% dan 98.66%, sementara *Naïve Bayes* hanya 61.38% dan 55.63%.

Untuk memvalidasi perbedaan ini, dilakukan Uji McNemar. Hasilnya (divisualisasikan pada Gambar 8) menunjukkan nilai p-*value* < 0.001, yang menyimpulkan adanya perbedaan yang signifikan secara statistik. Nilai p- yang sangat kecil ini mengindikasikan bahwa kemungkinan perbedaan kinerja kedua model terjadi secara kebetulan sangatlah rendah. Secara umum, p-*value* < 0.05 digunakan sebagai ambang batas untuk menolak hipotesis nol (tidak ada perbedaan), sehingga p-*value* < 0.001 memberikan tingkat keyakinan yang jauh lebih tinggi terhadap keberadaan perbedaan nyata. Kurva ROC pada Gambar 9 juga menegaskan superioritas model KNN, ditunjukkan oleh nilai *Area Under Curve* (AUC) yang lebih tinggi dibandingkan model *Naïve Bayes*.



Gambar 8. Visualisasi Tabel Kontingensi *McNemar's Test*



Gambar 9. Kurva ROC Model *Naïve Bayes* dan KNN pada Data Uji

Berdasarkan perbandingan metrik pada Tabel IX dan nilai *AUC* (Gambar 9), terlihat bahwa model KNN (K=3) secara konsisten menunjukkan kinerja yang jauh lebih unggul dalam memprediksi stunting pada dataset penelitian ini, dan perbedaan ini terbukti signifikan secara statistik melalui Uji McNemar ($p < 0.05$).

F. Pembahasan Hasil

Temuan utama dari penelitian ini adalah superioritas kinerja model *K-Nearest Neighbors* (KNN) dengan K=3 dibandingkan dengan model *Gaussian Naïve Bayes*, di mana keunggulannya terbukti signifikan secara statistik melalui Uji McNemar ($p < 0.05$). Kinerja *Naïve Bayes* yang relatif rendah (F1-*Score* Stunting: 55.63%) dapat dijelaskan oleh pelanggaran asumsi dasarnya. Analisis korelasi menunjukkan hubungan positif yang sangat kuat ($r \approx 0.87$) antara fitur 'Umur (bulan)' dan 'Tinggi Badan (cm)', yang mencederai asumsi independensi fitur *Naïve Bayes*. Hal ini didukung oleh kurva pembelajarannya yang menunjukkan pola *high bias*, mengindikasikan bahwa model terlalu sederhana untuk menangkap kompleksitas data.

Di sisi lain, keunggulan superior KNN (K=3) dengan F1-*Score* Stunting 98.66% dan *Recall* 99.25%, dapat diatribusikan pada sifatnya yang non-parametrik dan metodologi pra-pemrosesan yang cermat. Pemilihan K=3 merupakan keputusan kunci untuk mitigasi risiko *overfitting* yang teridentifikasi pada K=1 melalui analisis kurva validasi, menghasilkan model yang lebih robust dan dapat digeneralisasi. Hal ini didukung oleh kurva pembelajaran untuk KNN (K=3) yang menunjukkan pola *well-fitted model* (rendah bias dan rendah varians).

Secara metodologis, penelitian ini mengkonfirmasi pentingnya pra-pemrosesan data yang cermat (penskalaan, SMOTE) dan evaluasi menggunakan metrik yang relevan seperti *Recall* dan F1-*Score* untuk kelas minoritas. Namun, penelitian ini memiliki beberapa keterbatasan, antara lain lingkup algoritma yang hanya membandingkan *Naïve Bayes* dan KNN, keterbatasan fitur yang hanya menggunakan data antropometri dasar, cakupan data yang terbatas, dan belum adanya validasi eksternal.

IV. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan disimpulkan bahwa model prediksi stunting pada balita berhasil dibangun menggunakan algoritma *Naïve Bayes* dan *K-Nearest Neighbors* (KNN). Pembangunan model ini didukung oleh metodologi pra-pemrosesan data yang sistematis, meliputi eksklusi data, transformasi fitur (*Label Encoding* dan *StandardScaler*), serta penanganan data tidak seimbang pada data latih menggunakan teknik *oversampling* SMOTE, yang terbukti membantu dalam meningkatkan kemampuan model mengenali kasus stunting. Model *K-Nearest Neighbors* (KNN) dengan K=3 menunjukkan performa yang secara signifikan lebih unggul dibandingkan model *Naïve Bayes* dalam memprediksi status stunting pada *dataset* balita di Kecamatan Banjaran, dengan nilai Akurasi 99.11%, *Recall* (Stunting) 99.25%, dan F1-*Score* (Stunting) 98.66%. Algoritma KNN (K=3) menunjukkan hasil yang menjanjikan sebagai bukti konsep (*proof-of-concept*) untuk dikembangkan lebih lanjut. Namun, sebagai studi pendahuluan (*pilot study*) dengan keterbatasan pada penggunaan tiga fitur antropometri dasar (umur, jenis kelamin, dan tinggi badan) serta data yang hanya berasal dari satu Lokasi, validasi eksternal dan pengayaan fitur sangat diperlukan sebelum model ini dapat dipertimbangkan sebagai alat bantu yang andal dalam sistem deteksi dini stunting.

Saran untuk penelitian selanjutnya difokuskan pada tiga area utama. Pertama, pengayaan fitur dengan menambahkan variabel non-antropometri seperti status sosio-ekonomi dan riwayat gizi untuk membangun model yang lebih komprehensif. Kedua, validasi eksternal pada dataset dari wilayah lain sangat diperlukan untuk menguji kemampuan generalisasi model. Ketiga, melakukan perbandingan dengan algoritma yang lebih canggih, seperti *ensemble methods* (misalnya, *Random Forest*) serta eksplorasi teknik pra-pemrosesan dan *tuning hyperparameter* yang lebih lanjut.

DAFTAR PUSTAKA

[1] E. Alpaydin, *Introduction to machine learning*, 4th ed. MIT Press, 2020.

[2] C. M. Bishop and H. Bishop, *Deep Learning: Foundations and Concepts*, 1st ed. Springer, 2023. doi: 10.1007/978-3-031-45468-4.

[3] J. Han, M. Kamber, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*, 4th ed. Morgan Kaufmann, 2022.

[4] K. P. Murphy, *Probabilistic Machine Learning: An Introduction*. Cambridge, MA: MIT Press, 2022.

[5] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, *An Introduction to Statistical Learning: with Applications in Python*. Springer, 2023.

[6] Kementerian Kesehatan Republik Indonesia, *Buku Saku Hasil Survei Status Gizi Indonesia (SSGI) Tahun 2023*. Jakarta: Badan Kebijakan Pembangunan Kesehatan, 2023.

[7] Kementerian Kesehatan Republik Indonesia, *Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 tentang Standar Antropometri Anak*. Jakarta: Kemenkes RI, 2020.

[8] Kementerian Kesehatan Republik Indonesia, *Pedoman Pemantauan Pertumbuhan Anak*. Jakarta: Kemenkes RI, 2020.

[9] World Health Organization (WHO), *WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development*. Geneva: WHO Press, 2006.

[10] S. Almatsier, *Prinsip Dasar Ilmu Gizi*. Jakarta: Gramedia, 2011.

[11] A. Kukkar, R. Mohana, A. Kumar Singh, and P. Kumar, "Machine learning based quantitative approach for robust healthcare prediction system," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 12, pp. 6167–6183, 2020.

[12] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN Computer Science*, vol. 2, no. 3, p. 160, 2021, doi: 10.1007/s42979-021-00592-x.

[13] T. D. Pham, J. C. Ho, and H. Q. Nguyen, "Predicting stroke using machine learning algorithms," *Journal of Healthcare Engineering*, vol. 2020, p. 9151980, 2020, doi: 10.1155/2020/9151980.

[14] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A brief review of machine learning and ensembles based prediction for diabetes mellitus," *Multimedia Tools and Applications*, vol. 80, no. 28–29, pp. 35617–35667, 2021, doi: 10.1007/s11042-020-10348-x.

[15] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach for classification," in *Rough Sets: International Joint Conference, IJCRS 2020, Havana, Cuba, September 29 – October 2, 2020, Proceedings, Part II*, A. Bi, S. F. Qin, Y. Y. Yao, and J. T. Yao, Eds. Springer, 2020, pp. 986–997.

[16] D. Chicco, N. Tötsch, and G. Jurman, "The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation," *BioData Mining*, vol. 14, no. 1, p. 1, 2021, doi: 10.1186/s13040-021-00244-z.

[17] C. G. Walsh, B. N. Ribeiro, S. Dey, S. B. Kritchevsky, and D. B. Reuben, "Improving reporting standards for machine learning prognostic models in older adults: a multidisciplinary consensus statement," *Journal of the American Geriatrics Society*, vol. 69, no. 9, pp. 2672–2680, 2021, doi: 10.1111/jgs.17268.

[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[19] C. Fannany and P. H. Gunawan, "Analisis Klasifikasi Pembelajaran Mesin untuk Pencegahan Proaktif Stunting Anak di Bojongsoang: Sebuah Studi Komparatif," 2024.

[20] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Information Sciences*, vol. 513, pp. 429–441, 2020, doi: 10.1016/j.ins.2019.11.004.

[21] P. Vuttipittayamongkol, E. Elyan, and C. Jayne, "A review of data sampling and data augmentation for a deep learning model for medical image classification," *Electronics*, vol. 10, no. 3, p. 359, 2021, doi: 10.3390/electronics10030359.

[22] W. C. Wahyudin, F. M. Hana, dan A. Prihandono, "Prediksi Stunting pada Balita di Rumah Sakit Kota Semarang Menggunakan Naive Bayes," *Jurnal Ilmu Komputer dan Matematika*, hlm. 32–36, 2023.

[23] I. C. R. Drajana dan A. Bode, "Prediksi Status Penderita Stunting Pada Balita Provinsi Gorontalo Menggunakan K-Nearest Neighbor Berbasis Seleksi Fitur Chi Square," *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, vol. 5, no. 2, hlm. 163-170, 2022.

[24] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, 2020.

[25] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, and J. R. Foulds, *Data Mining: Practical Machine Learning Tools and Techniques*, 5th ed. Morgan Kaufmann, 2025.

[26] Z. H. Zhou, *Ensemble Methods: Foundations and Algorithms*, 2nd ed. Chapman and Hall/CRC, 2025.