

# Perbandingan NBC dan KNN dalam Klasifikasi Stunting pada Anak di Wilayah Pedesaan

## *Comparison of NBC and KNN in Classifying Stunting in Children in Rural Areas*

Betrisandi  
Program Studi Teknik Informatika  
Universitas Puhuwato  
Puhuwato, Indonesia  
betris.sin@email.com

Rahmat Thaib  
Program Studi Teknik Informatika  
Universitas Puhuwato  
Puhuwato, Indonesia  
rahmatthaib@email.com

Diterima : September 2025  
Disetujui : Januari 2026  
Dipublikasi : Januari 2026

**Abstrak**—Stunting merupakan salah satu permasalahan gizi kronis yang masih menjadi perhatian serius di Indonesia. Anak yang mengalami stunting tidak hanya mengalami hambatan dalam pertumbuhan fisik, tetapi juga mengalami gangguan perkembangan kognitif yang berpotensi menurunkan tingkat kecerdasan, prestasi belajar, hingga produktivitas di usia dewasa. Permasalahan dalam penelitian ini adalah tingginya prevalensi stunting pada anak di wilayah pedesaan. Tujuan penelitian ini untuk menganalisis performa Naïve Bayes Classifier (NBC) dan K-Nearest Neighbor (KNN) dan membandingkan kinerja kedua metode tersebut untuk menentukan metode yang paling optimal dalam klasifikasi status stunting pada anak. Metode Penelitian yang digunakan meliputi tahap pengumpulan data melalui observasi dan wawancara. Pengolahan dan analisis data dilakukan dengan membandingkan metode NBC dan KNN dalam klasifikasi stunting anak. Hasil penelitian ini menunjukkan bahwa metode NBC lebih tinggi akurasi yaitu 95,24% dan F1-score 97% jika dibandingkan dengan metode KNN yang akurasi sebesar 76,19% dan F1-score 86%. Jadi, metode NBC lebih optimal digunakan dalam klasifikasi stunting pada anak.

**Kata Kunci**—Stunting Anak; Klasifikasi; NBC; KNN; Perbandingan

**Abstract**—Stunting is one of the chronic nutritional problems that remains a serious concern in Indonesia. Children who experience stunting not only experience physical growth retardation, but also cognitive development disorders that have the potential to reduce intelligence, academic achievement, and productivity in adulthood. The problem in this study is the high prevalence of stunting in children in rural areas. The purpose of this study is to analyse the performance of the Naïve Bayes Classifier (NBC) and K-Nearest Neighbour (KNN) and compare the performance of the two methods to determine the most optimal method for classifying stunting status in children in accordance with the Research

*Master Plan with a focus on engineering and technology for improving ICT content and the research topic of big data technology development. The research methods used included data collection through observation and interviews. Data processing and analysis were carried out by comparing the NBC and KNN methods in classifying child stunting. The results of this study indicate that the NBC method has higher accuracy, namely 95.24% and an F1-score of 97%, compared to the KNN method, which has an accuracy of 76.19% and an F1-score of 86%. Therefore, the KNN method is more optimal for use in classifying stunting in children.*

**Keywords**— Child Stunting; Classification; NBC; KNN; Comparison

### I. PENDAHULUAN

Stunting merupakan masalah kesehatan kronis yang ditandai dengan tinggi badan anak yang lebih rendah dari standar usianya akibat kekurangan gizi dalam waktu yang lama. Penderita stunting sangat rentan terhadap penyakit, tingkat kecerdasan di bawah rata-rata dan pada akhirnya mempengaruhi produktivitas [1]. Terdapat banyak faktor penyebab masalah gizi pada anak-anak stunting, termasuk dari asupan gizi yang tidak tercukupi sejak masa bayi, hal ini bahkan bisa dimulai sejak dalam kandungan. Faktor lain seperti kondisi ekonomi keluarga dan penyakit atau infeksi pada masa bayi juga dapat memengaruhi asupan gizi anak [2].

Stunting merupakan masalah serius yang berdampak luas terhadap kualitas sumber daya manusia dan masa depan bangsa [3]. Dampak yang ditimbulkan akibat stunting dapat berupa jangka panjang dan dampak jangka pendek [4]. Stunting bukan merupakan keturunan sehingga dapat dicegah. Pemenuhan gizi yang optimal pada masa 1.000 HPK yang dimulai sejak ibu hamil sampai anak usia 2 tahun merupakan investasi dalam mencegah stunting, upaya yang dapat dilakukan dalam percepatan pencegahan stunting yaitu melalui invertensi gizi spesifik dan intervensi gizi sensitif [5].

Di Indonesia, stunting merupakan salah satu permasalahan gizi kronis yang masih menjadi perhatian

serius, khususnya di wilayah pedesaan yang memiliki akses terbatas terhadap fasilitas kesehatan, informasi gizi, serta sanitasi yang layak. Menurut BKKBN 2022 prevalensi stunting di pedesaan cenderung lebih tinggi dibandingkan dengan wilayah perkotaan. Prevalensi stunting di Indonesia mencapai 21,6% pada Tahun 2022. Berdasarkan data Survey Status Gizi Indonesia (SSGI) tahun 2022 prevalensi balita stunting di Provinsi Gorontalo mencapai 23,8% sehingga menempatkan Gorontalo menduduki peringkat 17 tertinggi secara nasional.

Pencegahan dini adalah salah satu cara terbaik untuk mengurangi prevalensi stunting bagi semua masyarakat [6]. Upaya untuk mendeteksi dan mengklasifikasikan status stunting secara dini sangat penting guna menghindari dampak jangka panjang seperti penurunan kemampuan kognitif dan produktivitas di masa dewasa. Untuk mengatasi tingginya prevalensi stunting pada anak di wilayah pedesaan, diperlukan sistem deteksi dini yang cepat dan akurat dan dapat diimplementasikan di daerah dengan keterbatasan sumber daya. Salah satu upaya yang dapat dilakukan adalah dengan memanfaatkan metode kecerdasan buatan, khususnya *machine learning* untuk klasifikasi stunting pada anak berdasarkan data antropometri dan faktor resiko lainnya [7].

Dua metode klasifikasi yang cukup populer dan banyak digunakan dalam penelitian Kesehatan adalah Naïve Bayes Classifier (NBC) dan K-Nearest Neighbor (KNN). NBC menggunakan pendekatan *probabilistic* berdasarkan teorema *bayes* mengasumsikan independensi antar fitur. Keunggulan utama metode ini adalah efisiensinya dalam komputasi serta kemampuannya menangani dataset besar dengan baik [7][8][9]. Disisi lain KNN adalah metode klasifikasi berbasis jarak yang menentukan kelas suatu data baru berdasarkan kedekatannya dengan sejumlah data tetangga terdekat [10][11].

Perbandingan antara NBC dan KNN dalam klasifikasi stunting menjadi penting karena masing – masing metode memiliki karakteristik dan keunggulan tersendiri tergantung pada struktur dan distribusi data. Beberapa penelitian menunjukkan hasil berbeda terkait akurasi dan waktu pemrosesan dari kedua algoritma ini, sehingga perlu studi lebih lanjut dalam konteks klasifikasi stunting pada anak untuk mengetahui metode mana yang paling optimal [12]. Sehingga identifikasi dan penanganan anak yang berisiko stunting bisa dilakukan lebih awal, lebih efisien dan tepat sasaran sesuai dengan strategi nasional untuk memperkuat upaya pencegahan yang lebih efektif.

Beberapa penelitian terkait diantaranya penelitian dengan judul klasifikasi data *mining* untuk penentuan stunting pada balita menggunakan metode *Naïve Bayes* dengan jumlah data 5.829 data balita menghasilkan akurasi yang sangat tinggi yaitu mencapai 95,08% [13]. Penelitian terkait lainnya dengan judul klasifikasi berisiko stunting pada balita dengan perbandingan KNN, *Naïve Bayes* dan SVM dengan beberapa variabel diantaranya umur, berat badan, panjang badan dengan jumlah data 523 data balita dengan rasio 80:20 data latih dan data uji hasil menunjukkan akurasi SVM 90%, KNN 89%, dan Naive Bayes 85% [14]. Dan penelitian terkait dengan judul penerapan algoritma *K-Nearest Neighbor* dalam klasifikasi status stunting balita di Desa Bojongemas dengan hasil akurasi sebesar 92% dengan jumlah 503 data balita [15].

Penelitian ini menekankan kebaruan melalui penggunaan data terbaru tahun 2025 serta penerapan variabel-variabel spesifik yang merepresentasikan kondisi aktual Kabupaten Pohuwato, yang belum banyak diakomodasi dalam penelitian sebelumnya. Sebagian besar studi perbandingan algoritma Naive Bayes Classifier (NBC) dan K-Nearest Neighbor (KNN) masih mengandalkan dataset lama atau data sekunder dengan karakteristik umum, sehingga kurang mencerminkan dinamika sosial dan geografis terkini di tingkat daerah. Data Pohuwato tahun 2025 memiliki tingkat heterogenitas yang tinggi, distribusi sampel yang tidak merata, serta dipengaruhi oleh faktor geografis dan sosial ekonomi lokal yang khas. Dengan perbedaan asumsi dan mekanisme kerja NBC yang berbasis probabilistik dan KNN yang bergantung pada kedekatan jarak antar data, penelitian ini mengkaji secara empiris kinerja kedua algoritma pada konteks data terbaru dan variabel lokal spesifik. Oleh karena itu, penelitian ini tidak hanya bersifat komparatif, tetapi juga memberikan kontribusi baru dalam pemilihan metode klasifikasi yang relevan untuk pengambilan keputusan berbasis data aktual di Kabupaten Pohuwato.

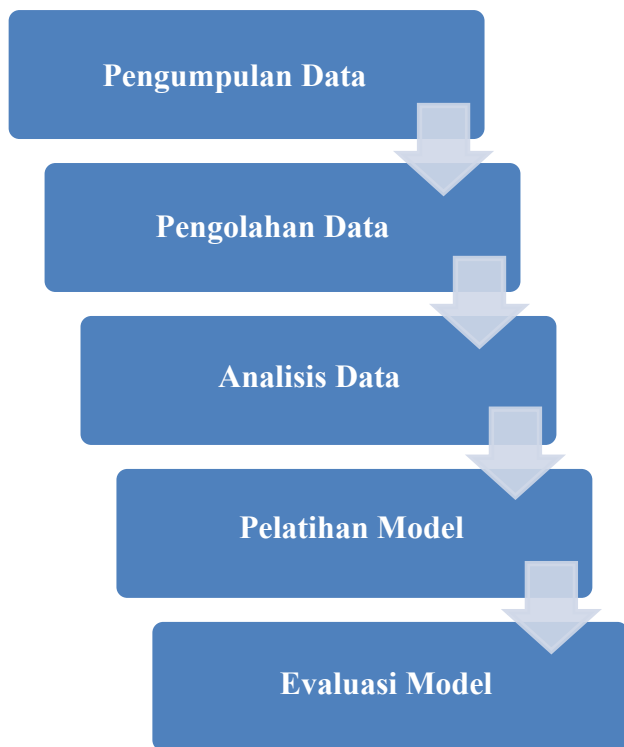
## II. METODE

Metode pada penelitian ini menggunakan metode kuantitatif dengan eksperimental komparatif. Tahapan-tahapan yang akan dilakukan terdiri dari tiga tahap. Tahap pertama, pengumpulan data dengan mengumpulkan data terkait dengan stunting pada anak yang diperoleh dari hasil observasi dan wawancara. Tahap kedua, analisis data dengan mengolah data stunting anak yang telah dikumpulkan kemudian akan diproses termasuk pembersihan data, normalisasi data dan transformasi. Langkah ini penting untuk mempersiapkan data yang baik sebelum digunakan untuk melatih model. Selanjutnya data akan dibagi menjadi dua subset yaitu data pelatihan (*training*) untuk melatih model klasifikasi dan data uji (*testing*) untuk menguji kinerja model. Selanjutnya proses latih model NBC dan KNN menggunakan data pelatihan. Model akan belajar pola-pola dari data pelatihan untuk mengklasifikasi stunting. Validasi model menggunakan data uji untuk metode NBC dan KNN, dilakukan untuk mengukur kinerja model dalam klasifikasi stunting pada anak dan terakhir evaluasi kinerja dari masing-masing metode NBC dan KNN menggunakan metrik evaluasi seperti akurasi, presisi, *recall* dan *F1-score* untuk membandingkan kinerja optimal dari kedua metode dalam klasifikasi serta menentukan metode yang paling efektif dalam mengklasifikasi stunting pada anak di pedesaan Kabupaten Pohuwato. Pemilihan nilai  $K=3$  pada K-Nearest Neighbor dilakukan dengan pertimbangan bahwa nilai  $K$  sangat memengaruhi performa klasifikasi. Nilai  $K$  yang terlalu kecil (misal  $K=1$ ) cenderung membuat model sensitif terhadap *noise* atau *outlier*, sehingga klasifikasi menjadi tidak stabil. Sebaliknya, nilai  $K$  yang terlalu besar (misal  $K>5$ ) dapat mengaburkan pola lokal spesifik dan menurunkan kemampuan KNN menangkap variasi unik data Kabupaten Pohuwato yang heterogen. Dengan  $K=3$ , model diharapkan menyeimbangkan sensitivitas terhadap pola lokal dan stabilitas terhadap outlier, sehingga memberikan akurasi yang lebih optimal pada dataset lokal yang relatif kecil dan tersebar tidak merata.

Tahap ketiga, penyusunan laporan dan publikasi yaitu dengan membuat laporan kemajuan, laporan akhir, dan

laporan publikasi dari penelitian yang akan dipublikasikan pada jurnal nasional terakreditasi SINTA 3 yaitu *Journal of electrical and Electronics Engineering (JAMBURA)*.

Tahapan penelitian lebih jelasnya dapat dilihat pada Gambar 1 berikut.



Gambar 1. Diagram Alir Penelitian

Strategi pemecahan masalah dalam penelitian ini yaitu mengumpulkan data anak usia balita dari wilayah pedesaan di Kabupaten Pohuwato yang mencakup indikator gizi dan faktor-faktor yang memengaruhi status stunting, *preprocessing* data dengan melakukan pembersihan, normalisasi, dan transformasi data agar siap digunakan dalam model klasifikasi, membangun dua model klasifikasi menggunakan metode NBC dan KNN untuk klasifikasi stunting dari data yang tersedia, mengevaluasi model dengan menggunakan metrik evaluasi, serta menganalisis hasil dan menyusun analisis dari hasil perbandingan, lalu memberikan rekomendasi terkait metode yang paling efektif dan efisien untuk diterapkan dalam sistem pendukung pengambilan keputusan di bidang Kesehatan masyarakat pedesaan. *Software* yang akan digunakan untuk memproses dataset pada penelitian ini yaitu rapidminer studio atau Altair AI Studio 2025.1.1.

### III. HASIL DAN PEMBAHASAN

#### A. Pengumpulan Data

Berdasarkan data dari Dinas Kesehatan Kabupaten Pohuwato, terdapat 3 kecamatan di Kabupaten Pohuwato yang tinggi angka status stunting pada anak yaitu Kecamatan Popayato terdiri dari 10 desa, Kecamatan Randangan terdiri dari 13 desa, dan Kecamatan Marisa terdiri dari 8 desa. Total data stunting pada anak dari 3 kecamatan tersebut berjumlah 102 data kasus stunting pada anak. Berikut merupakan contoh data stunting dari Dinas Kesehatan Kabupaten Pohuwato pada tabel 1.

TABEL 1. DATA STUNTING ANAK

No	JK	Tgl Lahir	Posyandu	Berat	Tinggi	BB/U	ZS BB/U	TB/U	ZS TB/U	BB/TB	ZS BB/TB	Naik Berat Badan
1	L	24/04/2022	SRIKANDI	10,5	89	Kurang	-2,78	Pendek	-2,2	Gizi Kurang	-2,4	T
2	L	08/01/2023	SRIKANDI	10,8	84	Normal	-1,72	Pendek	-2,44	Gizi Baik	-0,5	N
3	L	23/06/2024	PERTIWI	8,2	69	Normal	-1,39	Pendek	-2,66	Gizi Baik	0,01	T
4	L	04/03/2024	KASIH IBU	8,1	70,1	Kurang	-2,26	Sangat Pendek	-3,72	Gizi Baik	-0,5	T
5	P	12/12/2022	PERMATA IBU	10	84	Normal	-1,95	Pendek	-2,05	Gizi Baik	-1,1	T
6	P	10/02/2024	KASIH TINI	6,7	69,6	Sangat Kurang	-3,31	Sangat Pendek	-3,37	Gizi Kurang	-2,1	T
7	L	22/03/2024	BUKIT INDAH 1	7	69,3	Sangat Kurang	-3,4	Sangat Pendek	-3,78	Gizi Kurang	-2,1	T
8	L	31/12/2023	BUKIT INDAH 1	8,2	73,5	Kurang	-2,48	Pendek	-2,79	Gizi Baik	-1,6	T

Dari contoh data stunting anak pada tabel 1 terlihat beberapa kriteria stunting pada anak yaitu berat badan, tinggi badan, BB/U = Berat Badan/Umur, BB/TB = Berat Badan/Tinggi Badan dan TB/U = Tinggi Badan/Umur.

#### B. Analisis Data

Dalam proses analisis data terdapat beberapa tahapan yang perlu dilakukan yaitu:

- **Pemrosesan Data**

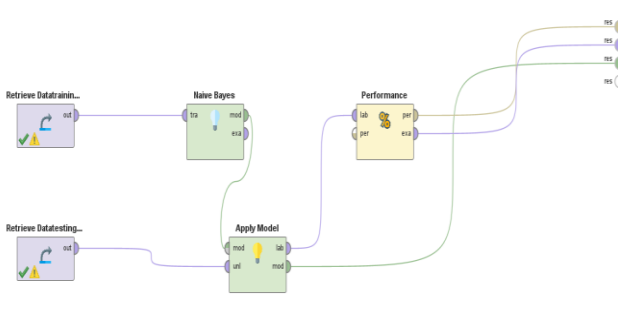
Pada tahap ini sebelum data diolah menggunakan aplikasi Rapidminer Studio, terlebih dahulu dilakukan pembersihan data dengan tujuan untuk mempersiapkan data yang digunakan dalam analisis sudah akurat. Proses ini dapat mendeteksi kesalahan data seperti data *double* atau data kosong serta menghapus data yang tidak sesuai kebutuhan. Dalam proses ini peneliti menghapus beberapa atribut data yang tidak diperlukan dalam olah data nantinya seperti nama anak, alamat, dan nama orang tua.

- **Pembagian Data**

Pembagian data dibagi menjadi 2 subset yaitu 80% data pelatihan (*training*) sebanyak 81 data untuk melatih model klasifikasi dan 20% data uji (*testing*) sebanyak 21 data untuk menguji kinerja model.

- **Pelatihan Model**

Tahapan selanjutnya proses latih model atau algoritma yang telah dipilih yaitu Naïve Bayes Classifier (NBC) dan K-Nearest Neighbor (KNN) dengan K=3 menggunakan data pelatihan sebanyak 81 data. Model akan belajar pola-pola data dari pelatihan untuk mengklasifikasi data stunting. Sedangkan validasi model menggunakan data uji sebanyak 21 data untuk mengukur kinerja model dalam klasifikasi stunting pada anak. Proses ini dilakukan pada aplikasi rapidminer studio atau Altair AI Studio 2025.1.1 seperti pada Gambar 2.



Gambar 2. Desain Model NBC

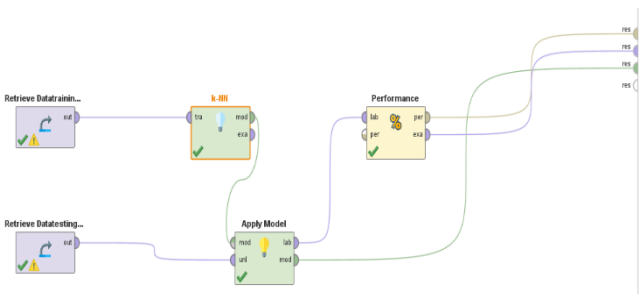
Gambar 2 menunjukkan desain model menggunakan NBC. Data training dihubungkan ke *Naive Bayes* dengan tujuan untuk mempelajari pola data untuk menghasilkan model sedangkan data testing dihubungkan ke *apply model*, *naive bayes* dihubungkan ke *apply model* untuk melihat hasil prediksi dari klasifikasi, semakin banyak data label pada data uji yang cocok dengan prediksi artinya semakin baik NBC dan semakin tinggi akurasi yang dihasilkan, contoh data hasil klasifikasi seperti pada Gambar 2. Selanjutnya *performance* untuk mengukur seberapa cerdas model NBC dan seberapa tinggi hasil akurasi, semakin tinggi hasil akurasi artinya sebagian besar data uji berhasil diklasifikasikan dengan benar oleh NBC.

RowNo	TBU	prediction	T...	confidence...	confidence...	JK	Tgl Lahir	Umr saat s...	Berat	Tinggi	BBU	ZS BBU	BBTB	ZS BBTB	ZS TBU	
1	Pendek	Pendek	0	1	1.000	P	Mar 3, 2022	3 tahun, 0 bul...	13	88	Normal	-0.860	Gro Baik	0.750	-2.390	
2	Pendek	Pendek	0	1	1.000	L	Apr 15, 2023	2 tahun, 4 bul...	12	82	Normal	-0.560	Gro Baik	0.910	-2.250	
3	Pendek	Pendek	0	1	1.000	L	Sep 8, 2024	1 tahun, 9 bul...	8	66	Normal	-1.530	Gro Baik	0.160	-2.840	
4	Sangat Pendek	Sangat Pendek	1	0	0.000	L	Jul 9, 2023	2 tahun, 2 bul...	9	75	Sangat Kurang	-3.060	Gro Baik	-1.370	-4.090	
5	Pendek	Pendek	0	1	1.000	L	Jun 9, 2025	0 tahun, 3 bul...	3	45	Normal	-1.900	Gro Baik	0.260	-2.580	
6	Pendek	Pendek	0	1	1.000	L	Jun 17, 2025	0 tahun, 2 bul...	3	46	Normal	-1.900	Gro Baik	-0.470	-2.050	
7	Sangat Pendek	Pendek	0	1	1.000	L	Jul 30, 2021	4 tahun, 1 bul...	8	86	Sangat Kurang	-3.540	Gro Baik	-1.930	-3.900	
8	Pendek	Pendek	0	1	0.403	0.597	P	Oct 27, 2020	4 tahun, 7 B...	1	95	Kurang	-2.410	Gro Baik	-1.200	-2.580
9	Pendek	Pendek	0	1	1.000	P	Oct 5, 2021	3 tahun, 7 B...	0	4	Kurang	-2.230	Gro Baik	-0.930	-2.640	
10	Pendek	Pendek	0	1	1.000	L	Oct 25, 2021	3 tahun, 7 B...	0	4	Sangat Kurang	-3.340	Gro Kurang	-2.570	-2.970	
11	Pendek	Pendek	0	1	1.000	L	Sep 8, 2022	2 tahun, 5 B...	11	82	Normal	-1.820	Gro Baik	-0.490	-2.820	
12	Pendek	Pendek	0	1	1.000	L	Aug 19, 2023	1 tahun, 5 B...	10	74	Normal	-0.580	Gro Baik	0.730	-2.390	
13	Pendek	Pendek	0	1	1.000	L	Aug 25, 2021	3 tahun, 5 B...	10	89	Sangat Kurang	-3.140	Gro Kurang	-2.520	-2.670	
14	Sangat Pendek	Pendek	0	1	1.000	L	Mar 1, 2021	3 tahun, 11...	11	87	Sangat Kurang	-3.120	Gro Baik	-1.270	-3.820	
15	Pendek	Pendek	0	1	1.000	P	Apr 1, 2021	3 tahun, 10...	10	89	Sangat Kurang	-3.290	Gro Kurang	-2.250	-2.900	
16	Sangat Pendek	Pendek	0	1	1.000	P	Nov 24, 2021	3 tahun, 2 B...	10	84	Kurang	-2.740	Gro Baik	-0.990	-3.280	
17	Pendek	Pendek	0	1	1.000	P	Feb 7, 2022	3 tahun, 9 B...	11	85	Kurang	-2.240	Gro Baik	-1.010	-2.560	
18	Sangat Pendek	Pendek	0	1	1.000	P	Feb 8, 2022	2 tahun, 11...	10	83	Kurang	-2.400	Gro Baik	-0.650	-3.160	
19	Pendek	Pendek	0	1	1.000	L	Dec 4, 2022	2 tahun, 2 B...	10	80	Kurang	-2.170	Gro Baik	-0.930	-2.800	
20	Pendek	Pendek	0	1	1.000	P	Mar 24, 2023	1 tahun, 10...	9	78	Normal	-1.520	Gro Baik	-0.430	-2.320	
21	Pendek	Pendek	0	1	1.000	L	May 2, 2024	0 tahun, 9 B...	8	68	Normal	-1.390	Gro Baik	-0.280	-2.080	

Gambar 3. Example Set (Apply Model) Naive Bayes

Model NBC mampu menghasilkan tingkat kepercayaan (*confidence*) yang tinggi pada mayoritas data uji, terbukti dengan nilai probabilitas prediksi yang dominan berada pada kisaran tinggi.

Selanjutnya lakukan kembali model dengan menggunakan metode KNN seperti pada Gambar 4 berikut.



Gambar 4. Desain Model KNN

Langkah-langkah desain model masih sama yang berbeda hanya pada metode yaitu metode KNN dengan  $K=3$ . Data *training* dihubungkan ke KNN dengan tujuan untuk mempelajari pola data untuk menghasilkan model sedangkan data *testing* dihubungkan ke *apply model*, KNN dihubungkan ke *apply model* untuk melihat hasil prediksi dari klasifikasi, semakin banyak data label pada data uji yang cocok dengan prediksi artinya semakin baik KNN dan semakin tinggi akurasi yang akan dihasilkan, contoh data hasil klasifikasi seperti pada Gambar 5. Selanjutnya *performance* untuk mengukur seberapa cerdas model KNN dan seberapa tinggi hasil akurasi.

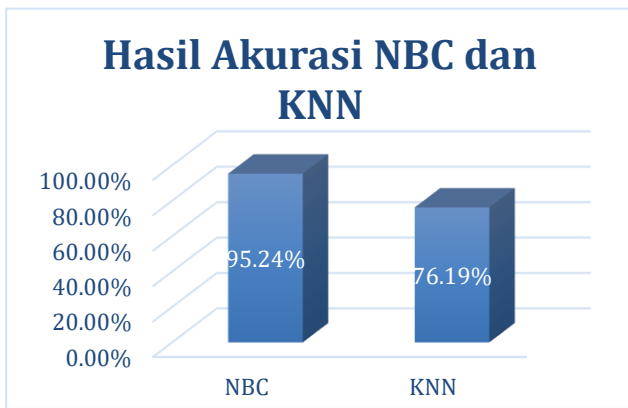
RowNo	TBU	prediction	T...	confidence...	confidence...	JK	Tgl Lahir	Umr saat s...	Berat	Tinggi	BBU	ZS BBU	BBTB	ZS BBTB	ZS TBU	
1	Pendek	Pendek	0	1	1.000	P	Mar 3, 2022	3 tahun, 0 bul...	13	88	Normal	-0.860	Gro Baik	0.750	-2.390	
2	Pendek	Pendek	0	1	1.000	L	Apr 15, 2023	2 tahun, 4 bul...	12	82	Normal	-0.560	Gro Baik	0.910	-2.250	
3	Pendek	Pendek	0	1	1.000	L	Sep 8, 2024	1 tahun, 9 bul...	8	66	Normal	-1.530	Gro Baik	0.160	-2.840	
4	Sangat Pendek	Pendek	0	1	1.000	L	Jul 9, 2023	2 tahun, 2 bul...	9	75	Sangat Kurang	-3.060	Gro Baik	-1.370	-4.090	
5	Pendek	Pendek	0	1	1.000	L	Jun 9, 2025	0 tahun, 3 bul...	3	45	Normal	-1.900	Gro Baik	0.260	-2.580	
6	Pendek	Pendek	0	1	1.000	L	Jun 17, 2025	0 tahun, 2 bul...	3	46	Normal	-1.900	Gro Baik	-0.470	-2.050	
7	Sangat Pendek	Pendek	0	1	1.000	L	Jul 30, 2021	4 tahun, 1 bul...	8	86	Sangat Kurang	-3.540	Gro Baik	-1.930	-3.900	
8	Pendek	Pendek	0	1	0.403	0.597	P	Oct 27, 2020	4 tahun, 7 B...	1	95	Kurang	-2.410	Gro Baik	-1.200	-2.580
9	Pendek	Pendek	0	1	1.000	P	Oct 5, 2021	3 tahun, 7 B...	0	4	Kurang	-2.230	Gro Baik	-0.930	-2.640	
10	Pendek	Pendek	0	1	1.000	L	Oct 25, 2021	3 tahun, 7 B...	0	4	Sangat Kurang	-3.340	Gro Kurang	-2.570	-2.970	
11	Pendek	Pendek	0	1	1.000	L	Sep 8, 2022	2 tahun, 5 B...	11	82	Normal	-1.820	Gro Baik	-0.490	-2.820	
12	Pendek	Pendek	0	1	1.000	L	Aug 19, 2023	1 tahun, 5 B...	10	74	Normal	-0.580	Gro Baik	0.730	-2.390	
13	Pendek	Pendek	0	1	1.000	L	Aug 25, 2021	3 tahun, 5 B...	10	89	Sangat Kurang	-3.140	Gro Kurang	-2.520	-2.670	
14	Sangat Pendek	Pendek	0	1	1.000	L	Mar 1, 2021	3 tahun, 11...	11	87	Sangat Kurang	-3.120	Gro Baik	-1.270	-3.820	
15	Pendek	Pendek	0	1	1.000	P	Apr 1, 2021	3 tahun, 10...	10	89	Sangat Kurang	-3.290	Gro Kurang	-2.250	-2.900	
16	Sangat Pendek	Pendek	0	1	1.000	P	Nov 24, 2021	3 tahun, 2 B...	10	84	Kurang	-2.740	Gro Baik	-0.990	-3.280	
17	Pendek	Pendek	0	1	1.000	P	Feb 7, 2022	3 tahun, 9 B...	11	85	Kurang	-2.240	Gro Baik	-1.010	-2.560	
18	Sangat Pendek	Pendek	0	1	1.000	P	Feb 8, 2022	2 tahun, 11...	10	83	Kurang	-2.400	Gro Baik	-0.650	-3.160	
19	Pendek	Pendek	0	1	1.000	L	Dec 4, 2022	2 tahun, 2 B...	10	80	Kurang	-2.170	Gro Baik	-0.930	-2.800	
20	Pendek	Pendek	0	1	1.000	P	Mar 24, 2023	1 tahun, 10...	9	78	Normal	-1.520	Gro Baik	-0.430	-2.320	
21	Pendek	Pendek	0	1	1.000	L	May 2, 2024	0 tahun, 9 B...	8	68	Normal	-1.390	Gro Baik	-0.280	-2.080	

Gambar 5. Example Set (Apply Model) KNN

KNN mengklasifikasi data berdasarkan kedekatan jarak antar data. Data Kabupaten Pohuwato bersifat heterogen dan tidak merata yaitu beberapa wilayah memiliki banyak sampel, sementara wilayah lain sangat sedikit. Ketidakeimbangan ini menyebabkan jarak ke tetangga terdekat ( $K=3$ ) tidak selalu merepresentasikan kelas sebenarnya, sehingga Performa KNN lebih rendah (76,19%) dibanding NBC (95,24%). Sedangkan NBC lebih stabil pada dataset kecil dan tidak seimbang karena memanfaatkan probabilitas fitur secara keseluruhan, sehingga lebih adaptif pada kondisi data lokal yang heterogen.

• Evaluasi Model

Tahapan terakhir melakukan evaluasi kinerja dari masing-masing metode yaitu NBC dan KNN menggunakan metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-Score*. *F1-Score* merupakan metrik evaluasi yang digunakan untuk mengukur keseimbangan antara *precision* (ketepatan) dan *recall* (cakupan) untuk membandingkan kinerja optimal dari kedua metode tersebut serta menentukan metode yang paling efektif dalam klasifikasi stunting pada anak. Hasil akurasi menunjukkan metode NBC adalah sebesar 95,24% sedangkan metode KNN adalah 76,19%. Lebih jelasnya dapat dilihat pada Gambar 6.



Gambar 6. Grafik Hasil Akurasi NBC dan KNN

Pada Gambar 6 terlihat jelas perbedaan hasil akurasi antara NBC dan KNN. Hal ini mengindikasikan gap performa yang signifikan sebesar 19,05% antara NBC dan KNN. Selanjutnya untuk menghitung nilai *precision*, *recall*, dan *F1-Score* menggunakan rumus:

$$\text{Precision (P)} : P = \frac{TP}{TP+FP}$$

$$\text{Recall (R)} : R = \frac{TP}{TP+FN}$$

Dimana :

- TP (True Positive) : jumlah data positif yang diprediksi benar
- FP (False Positive) : jumlah data negatif yang salah diprediksi positif
- FN (False Negative) : jumlah data positif yang salah diprediksi negatif

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

TABEL 2. CONFUSION MATRIX METODE NBC

	True Pendek	True sangat Pendek	Class Precision
Prediksi Pendek	15	0	100%
Prediksi Sangat Pendek	1	5	83,33%
Class recall	93,75%	100%	

Pada tabel 2, terlihat *precision* 100% untuk prediksi pendek dan 83,33% untuk prediksi sangat pendek. *Recall* 93,75% untuk *true* pendek dan 100% untuk *true* sangat pendek. Berdasarkan hasil *recall* diatas menunjukkan model ini sangat baik dalam mengklasifikasi kelas sangat pendek dengan *class Recall* 100%.

$$\text{Hitung nilai Precision (P)} : P = \frac{TP}{TP+FP} = \frac{15}{15+0} = \frac{15}{15} = 1 \text{ (100\%)}$$

$$\text{Hitung nilai Recall (R)} : R = \frac{TP}{TP+FN} = \frac{15}{15+1} = \frac{15}{16} = 0,9375 \text{ (93,75\%)}$$

$$\text{Hitung F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{1 \times 0,9375}{1 + 0,9375} = 0,97 \text{ (97\%)}$$

TABEL 3. CONFUSION MATRIX METODE KNN

	True Pendek	True sangat Pendek	Class Precision
Prediksi Pendek	16	5	76,19%
Prediksi Sangat Pendek	0	0	0,00%
Class recall	100%	0,00%	

Sedangkan pada tabel 3, terlihat *precision* 76,19% untuk prediksi pendek dan 0% untuk prediksi sangat pendek. *Recall* 100% untuk *true* pendek dan 0% untuk *true* sangat pendek. Berdasarkan tabel Confusion matrix diatas menunjukkan bahwa model tidak mampu mengenali pola ini sama sekali, klasifikasi dalam kelas sangat pendek dengan nilai *recall* dan *precision* 0% meskipun kelas lain masih dapat diklasifikasikan. Hal ini dikarenakan Persebaran data yang tidak merata, jumlah sampel kelas sangat pendek relatif sedikit dibanding kelas lain, sehingga tetangga terdekat (K=3) jarang berasal dari kelas tersebut. serta sensitivitas KNN terhadap pola lokal yang membuat kelas minoritas jarang mendapat representasi tetangga yang cukup. Hasil ini menegaskan keterbatasan KNN pada dataset lokal yang heterogen dan tidak seimbang, sekaligus menjelaskan mengapa NBC lebih stabil dalam kondisi serupa karena memanfaatkan probabilitas fitur secara keseluruhan.

$$\text{Hitung nilai Precision (P)} : P = \frac{TP}{TP+FP} = \frac{16}{16+5} = \frac{16}{21} = 0,7619 \text{ (76,19\%)}$$

$$\text{Hitung nilai Recall (R)} : R = \frac{TP}{TP+FN} = \frac{16}{16+0} = \frac{16}{16} = 1 \text{ (100\%)}$$

$$\text{Hitung F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{0,7619 \times 1}{0,7619 + 1} = 0,86 \text{ (86\%)}$$

TABEL 4. PERBANDINGAN AKURASI DAN F1-SCORE

	NBC	KNN
Akurasi	95,24%	76,19%
F1-Score	97%	86%

Berdasarkan tabel 4, hasil evaluasi dari kedua metode baik Akurasi dan *F1-Score* menunjukkan bahwa metode NBC lebih tinggi akurasinya yaitu 95,24% dan *F1-score* 97% jika dibandingkan dengan metode KNN akurasi sebesar 76,19% dan *F1-score* 86%. Dengan kesimpulan bahwa metode NBC lebih optimal digunakan dalam klasifikasi stunting pada anak.

#### IV. KESIMPULAN

Dari hasil perbandingan akurasi dan *F1-Score*, dapat dilihat bahwa metode NBC lebih tinggi akurasinya yaitu 95,24% dan *F1-score* 97% dibandingkan dengan metode KNN yang akurasinya hanya sebesar 76,19% dan *F1-score* 86%. Hasil ini menunjukkan bahwa metode NBC dapat digunakan dalam sistem pendukung pengambilan keputusan di bidang kesehatan masyarakat pedesaan. Manfaat penerapan metode ini yaitu dapat memudahkan tenaga kesehatan dalam deteksi dini penyakit stunting serta pencegahan stunting utamanya di daerah yang rawan kekurangan gizi. Saran untuk peneliti selanjutnya Untuk meningkatkan performa KNN, terutama pada kelas minoritas seperti *Sangat Pendek*, disarankan untuk menerapkan teknik oversampling seperti SMOTE (Synthetic Minority Oversampling Technique). Dengan SMOTE, jumlah sampel kelas minoritas dapat ditingkatkan secara sintesis sehingga model KNN memiliki representasi tetangga yang cukup, mengurangi bias terhadap kelas mayoritas, dan meningkatkan akurasi serta Recall pada kelas yang sebelumnya gagal diklasifikasikan.

## UCAPAN TERIMA KASIH

Terima kasih kepada Direktorat Riset, Teknologi, dan Pengabdian kepada Masyarakat (DRTPM) sebagai pemberi dana hibah pada Program Penelitian Dosen Pemula (PDP) tahun 2025.

## REFERENSI

- [1] Muchtar, Febriana, et al. "Edukasi pengenalan stunting pada remaja putri." *Lamahu: Jurnal Pengabdian Masyarakat Terintegrasi* 2.2 (2023): 138-144.
- [2] Pamuji, M. A. R., Prasetyo, M. B., & Kurniawan, T. R. (2023, January). Aplikasi Diagnosa Stunting Pada Balita Berbasis Android Menggunakan Metode Forward Chaining. In *Seminar Nasional Teknologi & Sains* (Vol. 2, No. 1, pp. 123-128).
- [3] Mulyana, A. (2025). Implikasi Stunting Terhadap Kemampuan Kognitif Dan Fisik Anak Usia Dini Di Tk At-Taqwa Teta Lambitu. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 10(02), 122-133..
- [4] Suryani, K., Rini, M. T., Hardika, B. D., & Widiastari, N. K. (2023). Analisis faktor penyebab kejadian stunting. *Jurnal Keperawatan Florence Nightingale*, 6(1), 8-12.
- [5] Sahroji, Q. N., Hidayat, R., & Nababan, R. (2022). Implementasi Kebijakan Dinas Kesehatan Dalam Penanganan Stunting Di Kabupaten Karawang. *Jurnal Pemerintahan Dan Politik*, 7(1).
- [6] Adistie, F., Lumbantobing, V. B. M., & Maryam, N. N. A. (2018). Pemberdayaan kader kesehatan dalam deteksi dini stunting dan stimulasi tumbuh kembang pada balita. *Media Karya Kesehatan*, 1(2).
- [7] Gustriansyah R, Suhandi N, Puspasari S, Sanmorino A. Machine Learning Method to Predict the Toddlers' Nutritional Status. *J Infotel*. 2024;16(1):32-43.
- [8] Jannah M. Perbandingan Metode Naïve Bayes Dan K-Nearest Neighbor Dalam Mengklasifikasi Status Pertumbuhan Anak Stunting (Studi Kasus: Posyandu Cemara). 2022;14(1):250-5. Available from: [https://repository.unilak.ac.id/id/eprint/2807%0Ahttps://repository.unilak.ac.id/2807/1/1855201057\\_BAB-I\\_VI\\_Daftar-Pustaka.pdf](https://repository.unilak.ac.id/id/eprint/2807%0Ahttps://repository.unilak.ac.id/2807/1/1855201057_BAB-I_VI_Daftar-Pustaka.pdf)
- [9] Turjo EA, Rahman MH. Assessing risk factors for malnutrition among women in Bangladesh and forecasting malnutrition using machine learning approaches. *BMC Nutr* [Internet]. 2024;10(1):1-25. Available from: <https://doi.org/10.1186/s40795-023-00808-8>.
- [10] Pangestu HG, Sinaga RY, Ulya FZ, Athiyah U, Muhammad AW, Alameka F. Analisis Efisiensi Metode K-Nearest Neighbor dan Forward Chaining Untuk Prediksi Stunting Pada Balita. *Inform Mulawarman J Ilm Ilmu Komput*. 2023;18(2):78.
- [11] Teknologi S, Ibrahimy U, Teknologi S, Ibrahimy U. Gudang Jurnal Multidisiplin Ilmu Implementasi Metode K-Nearest Neighbor ( K-NN ) Pada Klasifikasi Stunting Balita. 2024;2:282-8.
- [12] Loka SKP, Marsal A. Perbandingan Algoritma K-Nearest Neighbor dan Naïve Bayes Classifier untuk Klasifikasi Status Gizi Pada Balita. *MALCOM Indones J Mach Learn Comput Sci*. 2023;3(1):8-14.
- [13] Mulyanto Y, Idifitriani F, Wati A, Sumbawa UT, Mining D, Tano KP. Vol 7 No 2 , September 2024 Klasifikasi Data Mining Untuk Penentuan Stunting. 2024;7(2):119-25.
- [14] Dwinanto RW, A ASS, Ardianto R. Klasifikasi Berisiko Stunting Pada Balita: Perbandingan K- Nearest Neighbor , Naïve Bayes , Support Vector Machine. 2024;8(2):264-73.
- [15] Widia Pebrianti S, Astuti R, M Basysyar F. Penerapan Algoritma K-Nearest Neighbor Dalam Klasifikasi Status Stunting Balita Di Desa Bojongemas. *JATI (Jurnal Mhs Tek Inform*. 2024;8(2):2479-88.