

BAB III

METODOLOGI PENELITIAN

3.1 Metode Penelitian

Penelitian menggunakan pendekatan kuantitatif dengan metode klasifikasi prediktif untuk menganalisis data historis stunting balita. Model dikembangkan dengan mengombinasikan SMOTE (*Synthetic Minority Over-sampling Technique*) pada algoritma *Naïve Bayes Gaussian* untuk mengatasi ketidakseimbangan kelas dan meningkatkan akurasi prediksi status stunting.

Fokus penelitian adalah membangun model klasifikasi berbasis data sekunder yang mampu mengidentifikasi faktor-faktor dominan (seperti umur, berat badan, dan tinggi badan) yang memengaruhi stunting. Pendekatan ini dipilih karena kemampuan *Naïve Bayes Gaussian* dalam menangani data kontinu serta keunggulan SMOTE dalam menyeimbangkan distribusi kelas minoritas (kasus stunting).

3.2 Populasi dan Sampel

Populasi penelitian ini merupakan data sekunder yang diperoleh dari laman resmi Portal Jateng. Dataset tersebut berisi informasi mengenai status gizi dan stunting anak berdasarkan pengukuran berat badan (BB), tinggi badan (TB), lingkaran lengan atas (LiLA), serta indikator antropometri seperti BB/U, TB/U, dan BB/TB beserta *Z-Score*-nya. Data mencakup anak usia 0 tahun 4 bulan hingga 3 tahun 9 bulan dengan variasi kondisi gizi, mulai dari "Gizi Baik", "Gizi Kurang", hingga "Gizi Buruk". Atribut kunci meliputi Usia, parameter fisik (Berat, Tinggi, LiLA), klasifikasi status gizi (ZS BB/U, ZS TB/U, ZS BB/TB), serta indikator intervensi seperti pemberian PMT dan vitamin A.

No	Nama	JK	Usia Saat Ukur	Berat	Tinggi	LiLA	BB/U	ZS BB/U	TB/U	ZS TB/U	BB/TB
1.	ATIKA AZZAHRAH	P	1 Tahun - 8 Bulan - 26 Hari	8.5	75.5	0	Kurang	-2.03	Pendek	2.63	Gizi Baik
2.	ARSYAD BIL FAQIH	L	3 Tahun - 0 Bulan - 30 Hari	10.1	85.6	0	Sangat Kurang	-3.02	Pendek	2.97	Gizi Baik
3.	RAKA PRATAMA	L	1 Tahun - 0 Bulan - 8 Hari	6.3	70	0	Sangat Kurang	-3.82	Pendek	2.55	Gizi Buruk
4.	AISYAH SHIDQIYAH	P	3 Tahun - 9 Bulan - 5 Hari	9.9	90	0	Sangat Kurang	-3.51	Pendek	2.63	Gizi Kurang
	:		:		:		:		:		:
	:		:		:		:		:		:
	:		:		:		:		:		:
295.	AYRA KAMILA	P	1 Tahun - 6 Bulan - 19 Hari	8.9	75	14	Berat Badan Normal	-1.26	Pendek	2.16	Gizi Baik
296.	FASIKHA NUKHA KAMILA	P	4 Tahun - 3 Bulan - 13 Hari	10.5	91	15	Sangat Kurang	-3.48	Sangat Pendek	3.09	Gizi Kurang
297.	M KAFI ALFARUQ	L	1 Tahun - 2 Bulan - 18 Hari	8.4	73	16	Berat Badan Normal	-1.78	Pendek	2.28	Gizi Baik
298.	HUMAIRA AUFA RAHMA	P	2 Tahun - 11 Bulan - 0 Hari	11.1	85	0	Berat Badan Normal	-1.67	Pendek	2.49	Gizi Baik

Gambar 3. 1 Data Balita Stunting Kota Tegal 2024

Gambar 3.1 merupakan sampel penelitian ini merupakan bagian dari populasi data stunting Kota Tegal yang memenuhi kriteria kelengkapan informasi pada atribut utama, yaitu usia anak, berat badan (BB), tinggi badan (TB), lingkaran lengan atas (LiLA), serta nilai *Z-Score* untuk BB/U, TB/U, dan BB/TB. Data yang digunakan sebagai sampel telah melalui proses pembersihan dari duplikasi dan data tidak lengkap, sehingga hanya menyisakan entri yang valid untuk dianalisis.

3.3 Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian ini dengan metode dokumentasi, pengumpulan data yang bersumber dari dokumen atau arsip resmi yang telah ada. Data yang digunakan merupakan data sekunder yang diperoleh dari portal Jateng.

3.3.1 Dokumentasi

1. Mengunduh dataset langsung dari website Portal Jateng dalam format terstruktur (Excel/CSV).
2. Verifikasi data meliputi:
 - a. Pengecekan *missing values*.
 - b. Validasi konsistensi format (e.g., satuan berat/tinggi).

3.3.2 Studi Kepustakaan

1. Mengkaji literatur terkait:
 - a. Kriteria stunting menurut WHO (TB/U).
 - b. Penerapan Naïve Bayes Gaussian dan SMOTE dalam klasifikasi

kesehatan.

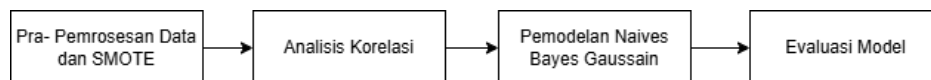
- c. Referensi jurnal terkait prediksi stunting berbasis machine learning.

3.3.3 Penyimpanan dan Pengolahan Awal

Data yang sudah siap kemudian disimpan dan diolah menggunakan Excel untuk proses visualisasi awal, lalu dilanjutkan ke tahap analisis lanjutan dengan menggunakan Python (*Pandas*, *NumPy*, *Scikit-learn*) dalam lingkungan Visual Studio Code.

3.4 Teknik Analisis Data

Analisis data penelitian ini melalui tahapan sistematis untuk memastikan akurasi model klasifikasi. Proses dimulai dengan pra-pemrosesan data, dilanjutkan dengan analisis korelasi, pemodelan, dan evaluasi performa. Dapat dilihat Gambar 3.2:



Gambar 3. 2 Tahapan Analisa Data

3.4.1 Pra-Pemrosesan Data

1. *Cleaning*:
 - a. Menghapus data duplikat untuk mencegah bias.
 - b. Mengatasi *missing values* dengan metode imputasi rata-rata (untuk data numerik).
 - c. Mengkonversi variabel Status ke bentuk numerik (contoh:, 0 = stunting, 1 = *Severely Stunted*).
2. Transformasi:
 - a. Melakukan normalisasi data numerik (Berat, Tinggi, Umur_Bulan) menggunakan *StandardScaler* agar skala seragam.
3. SMOTE:
 - a. Menerapkan teknik *oversampling* pada kelas minoritas (stunting) untuk menyeimbangkan distribusi dataset.

3.4.2 Analisis Korelasi

1. Membuat *heatmap* korelasi untuk mengidentifikasi hubungan antar variabel (misal: korelasi positif antara Tinggi dan Umur_Bulan).
2. Memvalidasi signifikansi variabel independen terhadap status stunting.

3.4.3 Pemodelan Naive Bayes Gaussian

1. Membagi dataset dengan rasio 80:20 (*training* dan *testing*).
2. Melatih model menggunakan fitur (Umur_Bulan, Berat, Tinggi) dan target (Status).
3. Mengoptimalkan parameter model dengan *cross-validation*.

3.4.4 Evaluasi Model

1. Mengukur performa model dengan metrik:
2. *Accuracy* (akurasi keseluruhan).
3. *Precision* dan *Recall* (kinerja pada kelas minoritas).
4. *F1-Score* (keseimbangan antara *precision* dan *recall*).
5. *Confusion Matrix* digunakan memvisualisasikan klasifikasi benar/salah pada tiap kelas.