BAB I
PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komputasi telah mendorong pengolahan data berskala besar menjadi semakin penting di berbagai bidang, seperti kesehatan, keuangan, dan industri. Salah satu pendekatan utama untuk memanfaatkan data ini adalah melalui *Machine Learning* (ML), yang memungkinkan komputer untuk belajar dari data tanpa pemrograman eksplisit. Algoritma ML bertujuan untuk menemukan pola atau tren dalam data yang dapat digunakan untuk membuat prediksi atau keputusan. Teknik ML yang sering digunakan adalah klasifikasi yang merupakan salah satu pendekatan paling banyak digunakan untuk memisahkan objek ke dalam kelas-kelas berdasarkan karakteristik tertentu (Rahmani *et al.*, 2021).

Dalam konteks kesehatan masyarakat, klasifikasi adalah salah satu teknik yang sering digunakan untuk mengelompokkan data menjadi kategori-kategori tertentu, misalnya untuk mengidentifikasi status gizi balita sebagai gizi baik atau gizi buruk (Hao *et al.*, 2020; Rahmani *et al.*, 2021). Klasifikasi dalam *supervised learning* adalah proses pengelompokan data *tren* dalam kategori yang telah diketahui berdasarkan label yang diberikan. Teknik tersebut sangat penting dalam berbagai aplikasi kecerdasan buatan dan statistik. Klasifikasi membangun model yang mampu memprediksi kelas atau kategori dari data baru berdasarkan variabel prediktor (Hao *et al.*, 2020). Beberapa algoritma yang umum digunakan untuk klasifikasi antara lain *Decision Trees* (Resmini *et al.*, 2021) *, K-Nearest Neighbors* (KNN) (Tempola *et al*., 2021), dan *Support Vector Machines* (SVM) (Géron, 2022). SVM adalah salah satu metode yang efektif dalam menangani masalah klasifikasi dengan memisahkan data berdasarkan variabel prediktor menggunakan *hyperplane* yang optimal untuk dua kategori data (Géron, 2022). Keunggulan SVM adalah kemampuannya untuk melakukan klasifikasi yang baik pada dataset yang kompleks, namun tetap efisien dalam hal generalisasi model (Bhavsar *et al.*, 2021).

Adapun penelitian-penelitian sebelumnya yang menggunakan SVM di antaranya Bleem & Virk (2020) yang menggunakan SVM untuk mendeteksi diabetes berdasarkan data kesehatan pasien dengan akurasi mencapai 77.92%. Ahmad & Polat (2023) melakukan kombinasi algoritma optimasi Jellyfish dan pengklasifikasi SVM memiliki performa tertinggi dengan akurasi yang diperoleh 94.48% dalam prediksi penyakit jantung. Selain itu, penelitian oleh Shanei *et al.* (2022) juga menggunakan SVM untuk mengklasifikasikan data non-linear dalam konteks pengenalan pola pada citra medis. Mereka mencapai akurasi 84% menggunakan kernel RBF, sementara penelitian oleh Resmini *et al.* (2021) berhasil mendiagnosis kanker payudara dengan akurasi 97,18% menggunakan kombinasi Algoritma Genetika (GA) dan SVM.

Penelitian-penelitian tersebut telah menunjukkan bahwa SVM sangat efektif dalam banyak kasus. Akan tetapi, ketika hubungan antara variabel prediktor dan respons tidak hanya bersifat non-linier tetapi juga halus atau bertahap, penggunaan kernel pada SVM menjadi kurang efektif. Oleh sebab itu, permasalahan ini dapat

diatasi dengan menggunakan fungsi spline, karena lebih fleksibel dan tidak mengasumsikan bentuk khusus dari hubungan antara variabel prediktor dan variabel respon (Kuyoro *et al.*, 2022; Tharwat, 2019).

Penggunaan fungsi spline dalam SVM memberikan fleksibilitas tambahan untuk memodelkan hubungan variabel prediktor dengan variabel respon tanpa asumsi linearitas (Kuyoro et al., 2022; Tharwat, 2019). Fungsi spline memungkinkan SVM untuk menangkap variasi data secara lebih detail, terutama di sekitar titik-titik simpul yang disebut *knots* (Helwig, 2020). Salah satu fungsi spline yang banyak digunakan dalam analisis adalah spline *truncated*, karena melibatkan titik *knot* dalam estimasi fungsinya. Titik *knot* merupakan titik perpaduan yang menandai perubahan perilaku fungsi pada interval yang berbeda. Pemilihan titik *knot* yang optimal dapat menentukan model spline terbaik. Salah satu metode untuk menentukan titik *knot* optimal adalah dengan mencari nilai *Generalized Cross Validation* (GCV) yang minimum (Islamiyati *et al.*, 2024).

Berdasarkan uraian sebelumnya, maka penelitian ini mengkaji metode klasifikasi SVM dengan fungsi spline kususnya spline *truncated* pada data status gizi balita di Kabupaten Gowa tahun 2022. Fokus utama penelitian adalah menerapkan metode SVM dengan pendekatan spline nonparametrik untuk menangkap pola hubungan yang kompleks dan bertahap antara rasio berat badan terhadap panjang badan (BB/PB) di Kabupaten Gowa. Data status gizi balita dikategorikan menjadi dua kelompok yaitu gizi buruk dan gizi baik. Selanjutnya data tersebut di klasifikan menggunakan SVM fungsi spline. Kinerja model dievaluasi berdasarkan nilai GCV dan persentase akurasi.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana penduga model klasifikasi SVM fungsi spline terbaik pada data status gizi balita di Kabupaten Gowa tahun 2022?
2. Bagaimana model klasifikasi SVM fungsi spline dalam mengklasifikasikan status gizi balita di Kabupaten Gowa tahun 2022?

1.3 Batasan Masalah

Penelitian ini hanya menggunakan data status gizi balita berdasarkan indikator berat badan menurut panjang badan (BB/PB) yang dikumpulkan dari Puskesmas di Kabupaten Gowa pada tahun 2022. Penelitian ini menggunakan metode SVM fungsi spline untuk mengklasifikasikan status gizi balita di Kabupaten Gowa tahun 2022, model yang digunakan dibatasi pada jumlah titik knot 1 hingga 3 serta orde yang digunakan linear, kuadratik, dan kubik. Data yang dianalisis mencakup variabel yang berhubungan dengan status gizi balita, seperti usia, berat badan, dan panjang/tinggi badan.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Memperoleh penduga model model SVM fungsi spline dalam mengklasifikasikan data status gizi balita di Kabupaten Gowa tahun 2022.
2. Memperoleh model klasifikasi SVM fungsi spline dalam mengklasifikasikan data status gizi balita di Kabupaten Gowa tahun 2022.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Mengembangkan metode SVM dengan fungsi spline dalam klasifikasi khususnya pada kasus status gizi balita di Kabupaten Gowa tahun 2022.
2. Meningkatkan akurasi klasifikasi status gizi balita, yang dapat mendukung program penilaian gizi balita di daerah-daerah termasuk di Kabupaten Gowa tahun 2022.

1.6 Teori

Sub-bab teori ini membahas mengenai beberapa teori yang digunakan untuk membahas pokok pembahasan terkait metode SVM fungsi Spline.

1.6.1 Fungsi spline *truncated*

Salah satu model regresi nonparametrik yang dapat digunakan untuk menduga kurva regresi adalah regresi spline. Regresi spline merupakan pendekatan yang mempertimbangkan kemulusan kurva dalam mencocokkan plot data. Kurva dibentuk berdasarkan titik-titik knot yang dihubungkan satu sama lain. Regresi spline memungkinkan penggunaan berbagai orde, sehingga dapat dibentuk regresi spline linier, kuadratik, kubik, maupun dengan orde lainnya. Spline adalah potongan-potongan polinomial (fungsi *truncated*) yang mulus namun tetap tersegmen (Wang, 2011). Penggunaan spline ini difokuskan pada pola data, di mana setiap daerah dapat memiliki karakteristik berbeda. Kemampuan ini ditunjukkan oleh fungsi truncated yang melekat pada estimator, dengan titik-titik potong polinomial yang disebut titik knot (Bintariningrum & Budiantara, 2014).

Pada spline truncated digunakan titik *knot. Knot* adalah nilai variabel prediktor yang menjadi batas antar segmen. *Knot* diartikan sebagai suatu titik fokus dalam fungsi spline, sehingga kurva yang dibentuk tersegmen pada titik tersebut (Dani & Ni'matuzzahroh, 2022). Fungsi spline dengan orde m dan titik-titik *knot* $K_1, K_2, \cdots, K_q$ yang didefinisikan sebagai sebarang fungsi f disajikan sebagai berikut:

$$f(x_i) = \sum_{k=0}^{m} \beta_k x_i^k + \sum_{p=1}^{q} \beta_{p+m} \left(x_i - K_p\right)_+^m \quad (1)$$

dan komponen *truncated*-nya:

$$\left(x_i - K_p\right)_+^m = \begin{cases} \left(x_i - K_p\right)^m & , \text{jika } x_i \geq K_p \\ 0 & , \text{jika } x_i < K_p \end{cases}$$

dengan:

x_i : nilai variabel prediktor ke-i

K_p : nilai titik *knot* ke-p

m : orde
β : nilai estimator

1.6.2 Pemilihan Titik *knot* Optimum

Penentuan model spline terbaik pada dasarnya melibatkan pemilihan titik *knot* yang optimal serta lokasi dari titik-titik *knot* tersebut. Proses ini juga mencakup pemilihan orde dari fungsi polinomial yang digunakan pada setiap segmen spline (Sari, 2016). Salah satu metode yang paling sering digunakan dalam memilih titik *knot* yang optimal adalah *Generalized Cross Validation* (GCV). GCV merupakan bentuk modifikasi dari metode *Cross Validation* (CV) yang digunakan untuk menilai kualitas model dalam konteks data (Islamiyati, 2019). Titik *knot* optimal dipilih berdasarkan nilai GCV yang paling minimum. Semakin kecil nilai GCV, semakin kecil pula galat dari model spline, yang menunjukkan model tersebut lebih representative Oleh karena itu, titik *knot* yang optimal akan memberikan model yang terbaik. Adapun persamaan GCV dapat dituliskan sebagai berikut (Wu & Zhang, 2006):

$$GCV = \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\left(1 - \frac{df}{n}\right)^2} \quad (2)$$

dengan:

y_i : nilai variabel respon ke-i
$\hat{y}_i$: nilai prediksi model ke-i
n : jumlah observasi
df : *degree of freedom* dari model spline

1.6.3 Klasifikasi

Klasifikasi merupakan suatu proses pencarian model atau fungsi yang membedakan konsep atau kelas data dengan tujuan untuk menebak kelas suatu objek yang tidak diketahui. Dalam klasifikasinya terdapat dua proses yaitu proses pelatihan (*training*) dan proses pengujian (*testing*). Proses pelatihan menggunakan *training set* yang telah diketahui labelnya yang berfungsi untuk membangun suatu model. Pengujian digunakan untuk menguji keakuratan model yang telah dibangun selama proses pelatihan (Resmini *et al.*, 2021).

1.6.4 SVM

SVM adalah salah satu metode yang sangat efektif dalam menangani masalah klasifikasi dengan memisahkan data menggunakan *hyperplane* yang optimal untuk dua kategori data (Géron, 2022). Fungsi pemetaan ini bisa berupa fungsi klasifikasi atau fungsi regresi. SVM menggunakan ruang *hyperplane* berupa fungsi-fungsi linear dalam *feature space*. *Support Vector Machines* dilatih menggunakan algoritma *supervised* yang didasarkan pada teori optimasi dengan menerapkan *learning bias* yang berasal dari teori pembelajaran statistik. Fungsi *hyperplane* didalam SVM dapat dilihat pada persamaan berikut (Ben-Hur *et al.*, 2008):

$$f(x_i) = wx_i + b \quad (3)$$

dengan:

$f(x_i)$: fungsi *hyperplane*
w : parameter *hyperplane*
x_i : variabel prediktor ke-*i*
b : bias

1.6.5 Klasifikasi SVM

Pengkalisifikasi SVM dengan mentukan margin dan *support vectors* terlebih dahulu. Margin adalah jarak antara *hyperplane* dan support vectors. *Support vectors* adalah titik data terdekat ke *hyperplane*. SVM bertujuan untuk memaksimalkan margin, yang artinya membuat pemisahan antar kelas sekuat mungkin. Adapun fungsi *hinge loss* SVM yang dituliskan pada persamaan berikut (Guerrero *et al.*, 2024):

$$\mathcal{L}(w,b) = \left|\left|\boldsymbol{w}\right|\right|^2 + \sum_{i=1}^{n} \max\left(0, 1 - y_i(wx_i + b)\right) \tag{4}$$

dengan:

y_i : nilai variabel respon ke-i, $y_i \in (-1, +1)$
w : parameter *hyperplane*
x_i : variabel prediktor ke-i
b : bias

Parameter w dan b tidak diketahui, sehingga dilakukan estimasi. Untuk mengestimasi parameter w dan b dengan cara melakukan iterasi dengan syarat $\min_{w,b} \frac{1}{2}\left|\left|w\right|\right|^2$ (Guerrero *et al.*, 2024). Selanjutnya, setelah nilai parameter w dan b diperoleh, klasifikasi pada data baru dapat dilakukan dengan menghitung:

$$\begin{aligned} \hat{y}_i &= sign(wx_i + b) \\ &= \begin{cases} +1 & \text{, jika } wx_i + b \geq 0 \\ -1 & \text{, jika } wx_i + b < 0 \end{cases} \end{aligned} \tag{5}$$

Jika hasilnya positif maka data diklasifikasikan ke kelas positif (+1) dan jika negatif data diklasifikasikan ke kelas negatif (-1) (Guerrero *et al.*, 2024).

1.6.6 Evaluasi Klasifikasi

Evaluasi dalam mengukur akurasi biasanya digunakan *confusion matrix*. *Confusion matrix* merupakan alat ukur berbentuk matrix yang digunakan untuk menghasilkan jumlah ketepatan klasifikasi terhadap *class* dengan algoritma yang digunakan (Haghighi *et al.*, 2018). Tebal *confusion matrix* dapat dilihat pada Tabel 1 berikut:

Tabel 1. *Confusion matrix*

Klasifikasi		**Prediksi**	
		Benar	**Salah**
Observasi	**Benar**	*True Positif (TP)*	*False Negatif (FN)*
	Salah	*False Positif (FP)*	*True Negatif (TN)*

Sumber (Haghighi *et al.*, 2018)

Keterangan:

True Positif (TP) : Jumlah sampel yang *true positif* yang di prediksi juga *true positif.*

True Negatif (TN) : Jumlah sampel yang *true negatif* yang di prediksi juga *true negatif.*

False Positif (TP) : Jumlah sampel yang *false positif* tetapi yang di prediksi *true positif.*

False Negatif (TN) : Jumlah sampel yang *false negatif* tetapi yang di prediksi *true negatif.*

Nilai *Accuracy* dapat dihitung menggunakan rumus berikut (Haghighi *et al.*, 2018):

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

Kemudian untuk melihat peluang kesalahan klasifikasi atau *Apparent Error Rate* (APER) dapat dihitung seperti pada Persamaan (6) berikut ini (Sulaehani, 2016):

$$APER = 1 - accuracy \tag{7}$$

1.6.7 *Stochastic Gradient Descent*

Stochastic Gradient Descent (SGD) merupakan metode optimasi yang populer untuk meminimalkan suatu fungsi, khususnya dalam model ML. Bentuk standar SGD dapat dijelaskan sebagai berikut (Liu *et al.*, 2020; Yu *et al.*, 2022):

1. Inisialisasi $w_0 = 0$ dan $b_0 = 0$
2. Iterasi SGD:
   Untuk setiap sampel data (x_i, y_i) dari dataset, akan diperbarui bobot *w* dan bias b sebagai berikut:

$$w_{i+1} = w_i - \eta \nabla_w \mathcal{L}(w, b) \tag{8}$$

$$b_{i+1} = b_i - \eta \nabla_b \mathcal{L}(w, b) \tag{9}$$

dengan:

η : *learning* rate

$\mathcal{L}(w, b)$: fungsi *hinge loss*

$\nabla_w \mathcal{L}(w, b)$: gradien *hinge loss* terhadap *w*

$\nabla_b \mathcal{L}(w, b)$: gradien *hinge loss* terhadap b

1.6.8 *Split data*

Split data adalah membagi dataset menjadi dua bagian yaitu data *training* dan data *testing* agar model bisa dilatih pada data *training* dan diuji pada data *testing* secara optimal (Uçar *et al.*, 2020). Tujuan utama membagi data adalah untuk memastikan bahwa model yang dibangun mampu menggeneralisasi dengan baik pada data baru,

bukan hanya sekadar mengingat data yang digunakan saat pelatihan. Hal ini menghindari overfitting, yaitu kondisi ketika model terlalu disesuaikan dengan data pelatihan dan tidak bekerja dengan baik pada data baru (Géron, 2022).

Salah satu metode *split data* yaitu *holdout validation. Holdout validation* adalah metode membagi dataset menjadi dua bagian (*training* dan *testing*) secara acak. Umumnya, 70%-80% dari data digunakan untuk *training* dan sisanya untuk *testing* (Tempola *et al*., 2021).

1.6.9 Status Gizi Balita

Status gizi adalah salah satu aspek dan indikator yang dapat menunjukkan pencapaian pembangunan kesehatan (Utami & Mubasyiroh, 2019). Hal ini dikarenakan aspek gizi berperan penting dalam pembentukan dan pembangunan sumber daya manusia. Salah satu fokus dalam intervensi dan masalah gizi yang masih terdapat di Indonesia maupun di dunia adalah gizi pada balita. Usia balita diklasifikasikan pada batasan nol hingga kurang dari lima tahun. Kelompok usia balita diketahui sebagai salah satu kelompok rentan gizi, berhubungan dengan masih tingginya masalah gizi kurang hingga gizi buruk, yang bedampak pada peningkatan untuk mengalami infeksi, penghambatan terhadap tumbuh kembang dan degradasi kondisi kesehatan di usia dewasa (Shanei *et al.*, 2022).

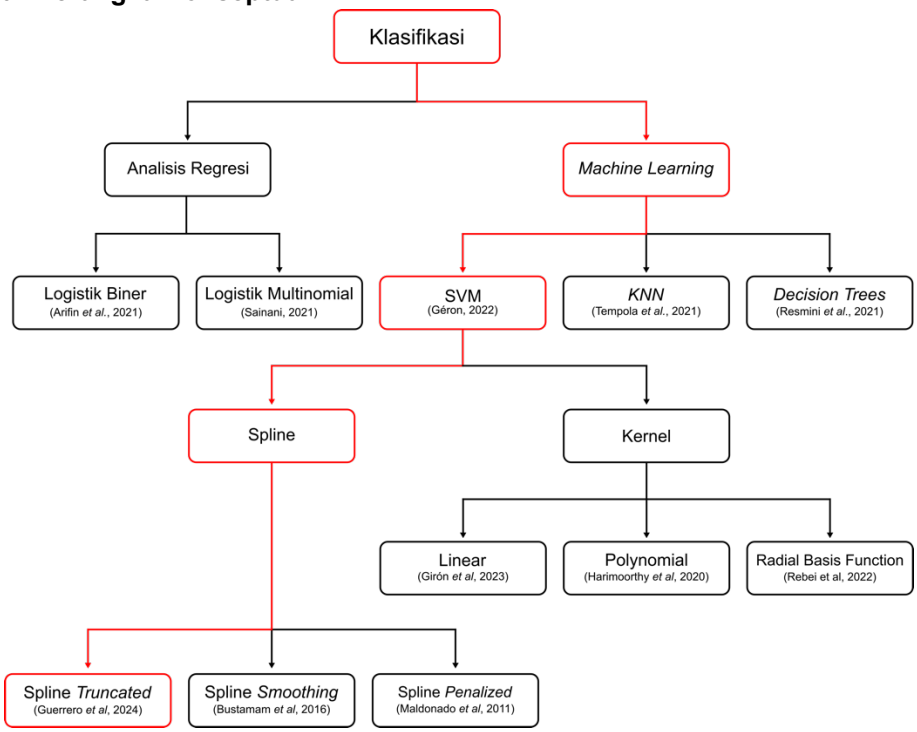
Gizi buruk diketahui sebagai salah satu permasalahan kesehatan yang belum tertangani dengan tuntas, sehingga diperlukan intervensi dan penanganan yang serius karena sifatnya yang iireversible atau tidak dapat kembali (Oktafiani *et al.*, 2024). Artinya, permasalahan gizi buruk dapat berdampak pada perkembangan balita yang terus berlangsung dalam jangka panjang, sehingga meningkatkan risiko morbiditas dan mortalitas. Salah satu bentuk gizi buruk adalah permasalahan stunting (pendek) dengan prevalensi sebesar 149 juta balita dan wasting (kerdil) dengan prevalensi sebesar 45 juta balita secara global pada 2020 (Fund, 2020). Fenomena gizi buruk dapat terjadi seiring dengan berbagai macam faktor yang melatarbelakangi timbulnya masalah gizi, baik dari faktor kesehatan, pendidikan, pengetahuan, kesadaran gizi, lingkungan, hingga asupan gizi yang diperoleh oleh balita. Oleh karena itu, diperlukan suatu upaya melalui program intervensi, edukasi dan promosi kesehatan untuk menurunkan risiko terjadinya gizi buruk pada balita (Lestari, 2022).

Pertumbuhan dan perkembangan balita dilakukandengan penilaian status gizi. Status gizi adalah keadaan keseimbangan antaraasupan zat gizi dari makanan dengan kebutuhan zat gizi yang diperlukan untuk metabolisme tubuh. Menurut Kementrian Kesehatan Republik Indonesia (Kemenkes RI) terdapat beberapa masalah status gizi pada balita yaitu *underweight* (gizi berat badan kurang), *stunted* (gizi pendek), dan *wasted* (gizi kurang). Gizi berat badan kurang adalah status gizi yang didasarkan pada berat badan menurut umur (BB/U), gizi pendek adalah status gizi yang didasarkan padapanjang badan menurut umur (PB/U) dan gizi kurang adalah status gizi yang didasarkan pada berat badan menurut tinggi badan (BB/PB)(Lestari, 2022).

Anak yang gizi buruk memiliki berat badan yang terlalu rendah untuk tinggi badan mereka, menunjukkan adanya masalah gizi akut atau infeksi yang mungkin dialami dalam waktu yang singkat. Gizi buruk dapat terjadi karena faktor-faktor seperti kurangnya asupan makanan, infeksi, atau kondisi lingkungan yang buruk yang memengaruhi status kesehatan balita (Lestari, 2022). Masalah gizi buruk sangat terkait dengan risiko kesehatan yang serius, termasuk risiko kematian yang lebih tinggi pada anak-anak yang mengalami gizi kurang. Maka dari itu, pentingnya klasifikasi dalam penilaian status gizi balita tidak bisa diabaikan. Dengan klasifikasi, tenaga kesehatan dapat mengidentifikasi jenis gizi buruk secara spesifik dan menentukan tingkat keparahannya. Klasifikasi memungkinkan pemberian penanganan yang tepat dan segera, sesuai dengan kondisi kesehatan anak. Hal ini sangat penting mengingat gizi buruk dapat berkembang menjadi masalah yang serius dalam waktu singkat, sehingga penanganan dini dapat mencegah komplikasi lebih lanjut yang berisiko fatal (Lestari, 2022).

Klasifikasi juga mendukung pelaksanaan intervensi gizi dan kebijakan kesehatan yang berbasis data. Dengan mengetahui prevalensi dan pola gizi buruk di suatu wilayah, kebijakan dapat disesuaikan untuk menjangkau kelompok balita yang paling rentan dan membutuhkan bantuan gizi. Ini mencakup intervensi nutrisi, edukasi keluarga, dan program pemantauan gizi yang intensif, yang bertujuan untuk memperbaiki status gizi balita secara komprehensif dan menurunkan angka kematian akibat gizi buruk (Utami & Mubasyiroh, 2019).

1.6.10 Kerangka Konseptual



Gambar 1. Kerangka Konseptual

BAB II
METODE PENELITIAN

2.1 Sumber Data

Penelitian ini menggunakan data sekunder berupa status gizi balita di Kabupaten Gowa pada tahun 2022. Data diperoleh dari Dinas Kesehatan Kabupaten Gowa dan merupakan hasil pengukuran status gizi balita berdasarkan berat badan menurut panjang badan (BB/PB) yang dilakukan di setiap posyandu di Kabupaten Gowa.

2.2 Variabel Penelitian

Variabel penelitian yang digunakan terdiri dari satu variabel respon (y) dan tiga variabel prediktor (x). Variabel respon dalam penelitian ini adalah status gizi balita berdasarkan berat badan menurut panjang badan (BB/PB) yang terdiri dari dua kategori yaitu gizi buruk dan gizi baik. Variabel prediktor merupakan faktor-faktor yang mempengaruhi status gizi balita seperti pada Tabel 2 berikut ini:

Tabel 2. Variabel Penelitian

Variabel Penelitian	Kode	Nama Variabel	Skala Pengukuran
Variabel Respon	y	Status Gizi Balita Berdasarkan Berat Badan Menurut Panjang Badan (BB/PB)	+1 = Gizi Buruk -1 = Gizi Baik
Variabel Prediktor	x_1	Usia (Bulan)	Rasio
	x_2	Berat Badan (Kg)	Rasio
	x_3	Panjang Badan (Cm)	Rasio

Sumber: Data olah, 2025

2.3 Metode Analisis Data

Adapun langkah-langkah yang dilakukan berdasarkan tujuan penelitian adalah sebagai berikut:

1. Estimasi parameter model klasifikasi SVM fungsi spline
2. Melakukan iterasi *Stochastic Gradient Descent* (SGD) untuk memperoleh estimator w dan b.
3. Membuat model klasifikasi SVM fungsi spline.
4. *Split data* menggunakan metode *holdout validation* dengan proporsi 80% data *training* dan 20% data *testing*.
5. Pemilihan titik *knot* optimum variabel prediktor usia (x_1) pada data *testing* untuk setiap orde dengan nilai GCV minimum.
6. Pemilihan titik *knot* optimum variabel prediktor berat badan (x_2) pada data *testing* untuk setiap orde dengan nilai GCV minimum.
7. Pemilihan titik *knot* optimum variabel prediktor panjang/tinggi badan (x_3) pada data *testing* untuk setiap orde dengan nilai GCV minimum.
8. Pemilihan model klasifikasi terbaik pada SVM fungsi spline
9. Pengujian model klasifikasi SVM fungsi spline pada data *testing*