

Metode K-Nearest Neighbor Dan Naive Bayes Dalam Menentukan Status Gizi Balita

Junius Pratama¹, Fauziah^{2*}, Ira Diana Sholihati³

^{1,2,3}Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi Dan Informatika, Universitas Nasional, Jakarta, Indonesia

E-mail: juniuspratama2018@student.unas.ac.id¹, fauziah@civitas.unas.ac.id², ira.diana@civitas.unas.ac.id³

Abstract

The nutritional intake needed by toddlers during the growth period for each individual has a different amount of consumption, therefore a process of checking nutritional status must be carried out. The indicators high and weight that will be calculated in determining the nutritional status. Until now the process of it status children under five is still carried out manually, resulting in the classification process of nutritional not being as expected. It is difficult for parents to go to the health center to check on their child's condition because the location is far from where they live and the administrative process is long. The propose is website-based information system, as well as providing recommendations on which method is the most accurate among K-Nearest Neighbors and Naïve Bayes in determining the nutritional status of toddlers. The results of this study have succeeded in creating an information system that can be used by users by inputting condition criteria for the process of classifying the nutritional status of toddlers. Naïve Bayes method is superior with a percentage value of 87.5% while the K-Nearest Neighbor method with $k = 3$ has an accuracy percentage of 71.25%.

Keywords: toddler nutrition, KNN, Naïve Bayes method.

Abstrak

Asupan gizi yang dibutuhkan balita pada saat masa pertumbuhan pada setiap individu memiliki jumlah konsumsi yang berbeda-beda, oleh sebab itu harus dilakukan proses pengecekan status gizi. Indikator dihitung dalam menentukan status gizi yaitu berdasarkan berat dan tinggi badan. Proses yang dilakukan dengan cara manual sehingga mengakibatkan proses klasifikasi status gizi kurang sesuai yang diharapkan. Orang tua sulit untuk mendatangi pusat kesehatan dengan mengecek kondisi anak karena lokasi yang jauh dari tempat tinggal serta proses administrasi yang lama. Tujuan penelitian ini adalah untuk mempermudah masyarakat khususnya orang tua dalam mengatasi permasalahan gizi yang dialami dengan cara memberikan fasilitas berupa sistem informasi berbasis website, serta memberikan rekomendasi metode mana yang paling akurat diantara K-Nearest Neighbor dan Naïve Bayes dalam menentukan status yang sesuai. Hasil dari penelitian ini telah berhasil membuat sebuah sistem informasi yang dapat digunakan oleh user dengan cara melakukan proses input kriteria kondisi untuk dilakukan proses klasifikasi status gizi. Tingkat akurasi metode Naïve Bayes lebih unggul dengan persentase nilai sebesar 87,5% sedangkan metode K-Nearest Neighbor dengan $k = 3$ mempunyai persentase akurasi sebesar 71,25%.

Kata Kunci: Gizi, Balita, K-Nearest Neighbor, Naïve bayes.

I. Pendahuluan

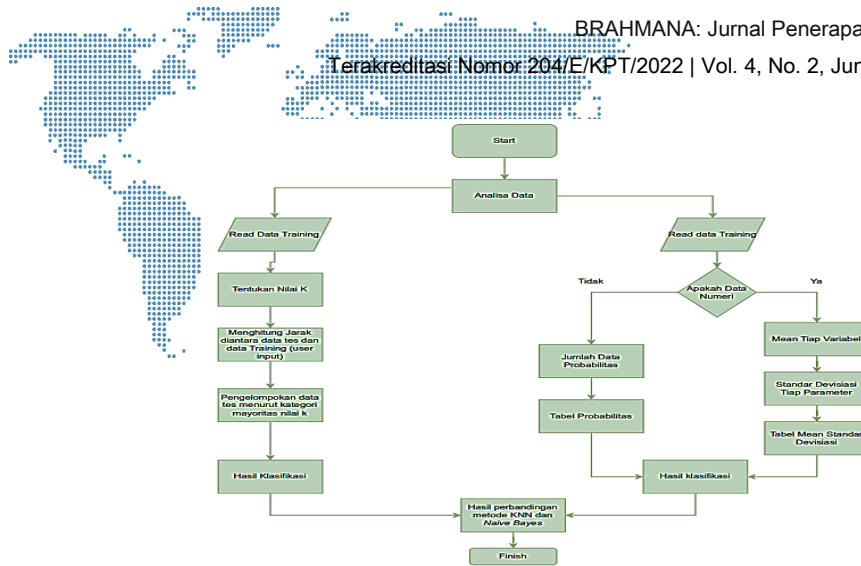
Setiap orang tua harus memperhatikan kebutuhan gizi anaknya untuk memaksimalkan proses tumbuh kembang. Selain itu kekurangan gizi dapat mengakibatkan masalah kesehatan dan kerusakan permanen[1]. Berdasarkan hasil Studi Kasus Gizi Indonesia pada tahun 2021 tingkat *stunted* pada balita sekitar 24,4%, tingkat *wasted* 7,1% dan

underweight sekitar 17,0%. *Stunted*, *wasted* dan *underweight* adalah tingkatan status gizi anak yang mana *stunted* adalah tingkat keadaan dimana kondisi pertumbuhan dan perkembangan anak mengalami gangguan, *wasted* kondisi ketika seorang balita memiliki berat badan rendah yang tidak sesuai dengan tinggi badannya sedangkan *underweight* adalah kondisi ketika anak memiliki berat badan terlalu rendah dibandingkan dengan usianya. Untuk mengetahui tingkat kesehatan pada balita dapat dilihat dari keadaan status gizi [2]. Penentuan status Gizi balitan dapat diukur dari beberapa indikator seperti tinggi badan, berat badan, dan berat badan menurut tinggi badan BB/TB[3]. Proses penentuan gizi dan pendataan terhadap balita diberbagai pusat kesehatan daerah di Indonesia masih banyak dilakukan secara manual menggunakan pembukuan tradisional. Berbagai aspek seperti pertimbangan dan kriteria harus diperhatikan untuk mempercepat proses penentuan status gizi balita. Metode ini menjadi kendala karena membutuhkan waktu yang lama dalam pendataan dan penentuan status gizi balita, oleh itu penulis akan mengambil masalah ini dan akan membangun suatu sistem yang dapat gunakan orang tua untuk mempermudah mengetahui status gizi anaknya. Berdasarkan permasalahan yang ada, dibutuhkan sebuah sistem penentu status gizi balita berbasis website untuk mempermudah dan mempercepat mendapatkan hasil yang benar. Sistem ini akan menggunakan dua metode data mining yaitu *K-Nearest neighbor* atau Naive Bayes. Algoritma *K-Nearest neighbor* digunakan untuk mengkalsifikasi data balita untuk menentukan status gizi balita berdasarkan jarak *Euclidean* terdekat, sedangkan mekanisme algoritma *Naive Bayes* mencari kemungkinan yang akan terjadi sebagai penentu keputusan. Penelitian dengan membandikan metode *K-Nearest Neighbor* dengan *Naive Bayes* untuk menentukan status gizi anak dibawah lima tahun. Nilai k dalam metode *K-Nearest Neighbor* mempunyai fungsi krusial sehingga dapat mempengaruhi hasil. Hasil dari penelitian ini menunjukkan bahwa metode *K-Nearest Neighbor* lebih unggul dari pada *Naive Bayes* dengan skor *if* lebih besar [4]. Penelitian yang ada kaitannya dengan proses klasifikasi status gizi menggunakan metode *K-Nearest Neighbor*. Hasil dari penelitian ini membuktikan bahwa metode *K-Nearest Neighbor* dapat digunakan untuk menentukan status gizi anak dibawah lima tahun. Skor persentase yang didapat mencapai 93.1% menggunakan nilai $k = 7$ [5]. Penelitian berikutnya membuat sebuah sistem *embedded* yang berguna untuk mengklasifikasi status gizi balita. Sistem ini mengimplementasikan metode *K-Nearest Neighbor* untuk mengklasifikasi data balita. Hasil pada saat proses pengujian sistem didapatkan hasil dengan persentase akurasi 97,14% [6]. Penelitian pada penentuan status gizi balita dengan mengimplementasikan metode *Naive Bayes*. Hasil dari penelitian ini menghasilkan tingkat akurasi dari 300 data dengan nilai sebesar 88% [7]. Penelitian kelima algoritma *Naive Bayes* digunakan untuk membuat sistem pakar sebagai pendeteksi status gizi balita berdasarkan perhitungan antropometri. Hasil penelitian ini menunjukkan tingkat kinerja dengan hasil 0.0013168617 [8]. Penggunaan dua metode ini diharapkan akan dapat meningkatkan tingkat akurasi dari hasil. Fokus penelitian ini pada pusat kesehatan masyarakat di daerah tidak dapat memberikan informasi mengenai status gizi balita secara efektif dan efisien karena proses pendataan dan penentuan gizi masih menggunakan cara manual dengan metode pembukuan secara tradisional. Hasil dari penentuan status gizi balita bahkan memiliki tingkat akurasi yang rendah, penyebabnya adalah pusat kesehatan daerah masih menggunakan metode lama atau data yang tidak update juga tidak mempunyai sistem untuk mengolah data yang ada. Tujuan Penelitian ini adalah mendesain sistem informasi berbasis website menggunakan bahasa pemrograman HTML, PHP dan basis data Mysql. Serta menerapkan metode *K-Nearest Neighbor* dan *Naive Bayes* pada sistem informasi status gizi balita.

2. Metodologi Penelitian

2.1. Analisa Desain Model

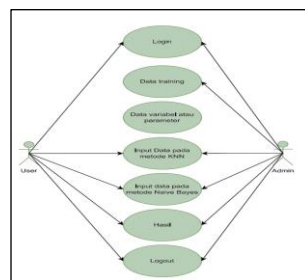
Tahapan-tahapan dari pemodelan metode *K-Nearest Neighbor* dan *Naive Bayes*.



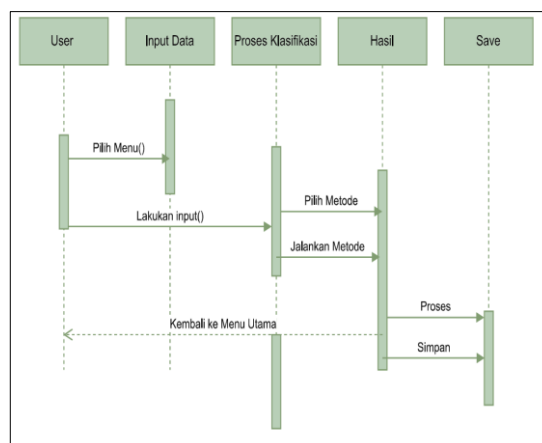
Gambar 1. Flowchart Analisa Perancangan Model

Dari flowchart perancangan model yang dijelaskan di Gambar 1. Dapat dideskripsikan analisa perancangan metode *K-Nearest Neighbor* dan *Naïve*

2.2. Analisa Desain Sistem



Gambar 2. Use Case Diagram



Gambar 3. Sequence Diagram

Dapat dijelaskan alur login ke sistem oleh user dengan menginputkan data kedalam sistem informasi penentuan status gizi balita. User dan admin mempunyai hak akses yang berbeda. User hanya dapat mengakses fasilitas seperti menggunakan metode untuk menentukan kondisi status gizi balita. Sedangkan admin dapat melakukan manipulasi data training seperti menambah , menghapus dan merubah data. Admin juga dapat melakukan manipulasi terhadap kriteria yaitu variabel atau parameter pada data training. Terdapat

empat parameter variable yang akan dimasukan yaitu umur atau usia anak, tinggi badan, Lingkar kepala dan berat badan. Selanjutnya algoritma akan memproses variabel yang di inputkan untuk kemudian melakukan prediksi hasil yang akan diklasifikasikan.

3. Hasil dan Pembahasan

3.1. Implementasi Sistem



Gambar 4. Tampilan Home

a. Menu Alternatif

Merupakan menu administrator management data training seperti menambahkan data baru, mengupdate isi data ataupun menghapus data. Menu alternatif merupakan tempat untuk melakukan proses training agar dapat memproses data testing.

ID	Nama Alternatif	Training	Opil
1	d1	Training	Usah Hapus
2	d2	Training	Usah Hapus
3	d3	Training	Usah Hapus
4	d4	Training	Usah Hapus
5	d5	Training	Usah Hapus
6	d7	Training	Usah Hapus

Gambar 5. Tampilan Alternatif

b. Training Data

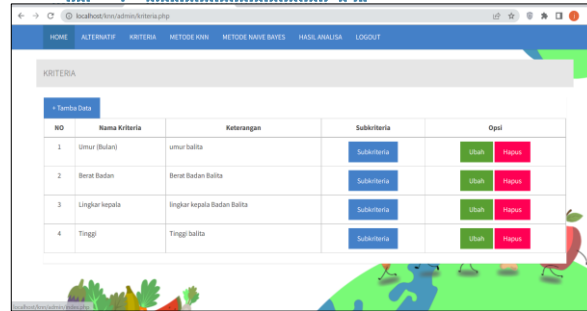
Merupakan menu untuk mentraining setiap data pada setiap atribut dan label atau klasifikasi.

ID	Nama alternatif	Umur (bulan)	Berat Badan	Lingkar kepala	Tinggi	Keputusan	Opil
1	d1	31	14.8	48	96	NORMAL	Usah Hapus

Gambar 6. Tampilan Data Training

c. Tampilan Menu Kriteria

Aplikasi ini didesain dapat digunakan untuk setiap kasus untuk metode *K-Nearest Neighbor* yang berbeda. Oleh sebab itu disediakan menu Kriteria. Menu kriteria merupakan fasilitas untuk membuat atribut data training berdasarkan kasus yang akan diselesaikan.



Gambar 7. Tampilan Kriteria

d. Tampilan Metode *K-Nearest Neighbor*

Merupakan menu untuk menampilkan kriteria. Tampilan Form Input Data Test. User dapat memasukkan Kriteria untuk menentukan status gizi bayi di form input kriteria lalu menekan tombol proses untuk memulai perhitungan.

Gambar 8. Tampilan *Form input*

e. Tampilan Data Training dan Data testing

Menampilkan data sebagai acuan metode dan mendapatkan output.

Data Training						
NO	Nama alternatif	Umur (Bulan)	Berat Badan	Lingkar kepala	Tinggi	Keputusan
1	d1	31	14.8	48	96	NORMAL
2	d2	46	14.1	50	95	NORMAL
3	d3	14	8.8	46	70	NORMAL
4	d4	51	14.2	49	94	NORMAL
5	d5	49	15.6	52	103	NORMAL
6	d7	21	11.9	46	89	NORMAL
7	d8	23	11.1	45	81	NORMAL
8	d9	6	8	42	61	NORMAL
9	d10	34	12.8	49	88	NORMAL
10	d11	12	8.2	45	71	NORMAL
11	d12	17	7.4	46	75	GIZI BURUK
12	d13	31	10.1	47	83	NORMAL
13	d14	24	12.5	47	82	NORMAL

Gambar 9. Tampilan Data *training* dan *testing*

f. Tampilan Perhitungan *Distance*

Hasil perhitungan jarak *Euclidean* data training dan data test.

Euclidean Distance						
NO	Nama alternatif	Umur (Bulan)	Berat Badan	Lingkar kepala	Tinggi	Distance
1	d1	676	96.04	1849	8281	104.41
2	d2	2116	198.81	2500	9025	117.64
3	d3	196	77.44	2116	4900	85.38
4	d4	2601	201.64	2401	8836	118.49
5	d5	2401	243.36	2704	10809	126.32
6	d7	441	141.61	2116	7921	103.05
7	d8	529	123.21	2025	6561	96.12
8	d9	36	64	1764	3721	74.73
9	d10	1156	163.84	2401	7744	107.07
10	d11	144	67.24	2025	5941	85.31
11	d12	289	54.76	2116	5625	89.92
12	d13	961	102.01	2209	6889	100.80
13	d14	576	156.25	2209	8724	96.31

Gambar 10. Tampilan *Euclidean Distance*

g. Tampilan Klasifikasi Berdasarkan *Minimum Distance*

Proses klasifikasi data menurut jarak *Euclidean* terdekat. Dan mayoritas klasifikasi $k = 3$.

Kode	Nama alternatif	Distance	Rangking	Pilihan	Keputusan
A14	d15	68.882	1	Ya	NORMAL
A15	d6	73.509	2	Ya	GIZI BURUK
A08	d9	74.733	3	Ya	NORMAL
A10	d11	85.307	4	Tidak	NORMAL
A03	d3	85.378	5	Tidak	NORMAL
A11	d12	89.915	6	Tidak	GIZI BURUK
A07	d8	96.116	7	Tidak	NORMAL
A13	d14	98.312	8	Tidak	NORMAL
A12	d13	100.802	9	Tidak	NORMAL
A06	d7	103.051	10	Tidak	NORMAL
A01	d1	104.413	11	Tidak	NORMAL
A09	d10	107.074	12	Tidak	NORMAL
A02	d2	117.443	13	Tidak	NORMAL

Gambar 11. Tampilan Klasifikasi *Euclidean Distance*

h. Tampilan Hasil

Output yang dihasilkan dari proses metode *K-Nearest Neighbor* penentuan status gizi balita.

Kesimpulan

Hasil perhitungan ini mengambil 3 data terbaik atau ascending ($K=3$) yang menggunakan **Klasifikasi Nearest neighbor**. Adapun Kesimpulan dari metode adalah **Kategori mayoritas (NORMAL) lebih banyak daripada mayoritas (GIZI BURUK)**, NORMAL berjumlah(2) sedangkan GIZI BURUK berjumlah 1 sehingga dapat disimpulkan atas nama **test keputusan kelayakan sebagai bantuan hasilnya : NORMAL**

Simpan hasil analisa

Gambar 12. Tampilan Hasil Keputusan

i. Tampilan Form Input Data Test

User dapat memasukan Kriteria untuk menentukan status gizi bayi di form input kriteria lalu menekan tombol proses untuk memulai perhitungan.

HOME ALTERNATIF KRITERIA METODE KNN METODE NAIVE BAYES HASIL ANALISA LOGIN/OUT

METODE Naive Bayes Classifier

Umur (bulan)

5

Berat Badan (kg)

5.8

Lingkar Kepala (cm)

40

Tinggi (cm)

54

PROSES

Gambar 13. Tampilan *form input* metode *Naive Bayes*

3.2. Tahapan menghitung probabilitas atribut data training.

PROBABILITAS KEPUTUSAN				
LAYAK	Jumlah Kejadian		Probabilitas	
	NORMAL	GIZI BURUK	NORMAL	GIZI BURUK
Probabilitas Keputusan	34	2	0.9750	0.0250

PROBABILITAS UMUR				
UMUR	Jumlah Kejadian		Probabilitas	
	NORMAL	GIZI BURUK	NORMAL	GIZI BURUK
31	2	0	0.14285714285714	0
40	1	0	0.071428571428571	0
14	1	0	0.071428571428571	0
51	1	0	0.071428571428571	0
48	1	0	0.071428571428571	0
21	2	0	0.14285714285714	0
23	1	0	0.071428571428571	0
6	1	0	0.071428571428571	0
24	1	0	0.071428571428571	0
12	1	0	0.071428571428571	0
17	0	1	0	0.5
24	1	0	0.071428571428571	0
5	1	0	0.071428571428571	0
7	0	1	0	0.5
Jumlah	14	1	1	1

Gambar 14. Tampilan Probabilitas

Tabel 1. Data Training

No	Umur (Bulan)	Berat Badan (Kg)	Lingkar Kepala (Cm)	Tinggi (cm)	Keputusan
1	31	14.8	48	96	NORMAL

No	Umur (Bulan)	Berat Badan (Kg)	Lingkar Kepala (Cm)	Tinggi (cm)	Keputusan
2	46	14.1	50	95	NORMAL
3	14	8.8	46	70	NORMAL
4	51	14.2	49	94	NORMAL
5	49	15.6	52	103	NORMAL
6	7	5.8	40	61	GIZI BURUK
7	21	11.9	46	89	NORMAL
8	23	11.1	45	81	NORMAL
9	6	8	42	61	NORMAL
10	34	12.8	49	88	NORMAL
11	12	8.2	45	71	NORMAL
12	17	7.4	46	75	GIZI BURUK
13	31	10.1	47	83	NORMAL
14	24	12.5	47	82	NORMAL
15	5	6.3	42	54	NORMAL

Tabel 2. Data Test

No	Umur (Bulan)	Berat Badan (Kg)	Lingkar Kepala (Cm)	Tinggi (cm)
1	5	5.8	40	54
2	14	8.8	46	70
3	31	10.1	47	83
4	17	7.4	46	75
5	7	5.8	42	61

Tabel 3. Jarak Minimum *Euclidean* data training dengan data test

No	Umur (Bulan)	Berat Badan (kg)	Lingkar kepala (cm)	Tinggi Badan (cm)	Keputusan	Jarak (Distance)
15	5	6.3	42	54	NORMAL	2.06155
6	7	5.8	40	61	GIZI BURUK	7.28011
9	6	8	42	61	NORMAL	7.67072
11	12	8.2	45	71	NORMAL	19.2031
3	14	8.8	46	70	NORMAL	19.5448
12	17	7.4	46	75	GIZI BURUK	24.9712
8	23	11.1	45	81	NORMAL	33.2579
14	24	12.5	47	82	NORMAL	35.1979
7	21	11.9	46	89	NORMAL	39.4235
13	31	10.1	47	83	NORMAL	39.8057
10	34	12.8	49	88	NORMAL	46.1194
2	46	14.1	50	95	NORMAL	59.4213
4	51	14.2	49	94	NORMAL	62.1897
1	31	14.8	48	96	NORMAL	62.6817
5	49	15.6	52	103	NORMAL	67.597

Tabel 4. Pengujian

Jumlah Data Training	Jumlah Data Testing	<i>K-Nearest Neighbor</i>	<i>Naïve Bayes</i>
15	5	60%	80%
100	40	75%	82,5%
150	60	71,66%	83,33%
200	80	71,25%	87,5%

Merupakan hasil pengujian yang didapatkan untuk proses pelatihan data dan direkomendasikan menggunakan *Naïve Bayes* pada penelitian ini

4. Kesimpulan

Dapat diambil kesimpulan bahwa Sistem informasi berbasis website yang dapat memudahkan masyarakat untuk melakukan pengecekan kondisi status gizi pada anak mereka. Penggunaan sistem informasi ini juga sangat sederhana *user* hanya memasukan data umur, berat badan, lingkar kepala dan tinggi lalu menekan tombol proses lalu sistem

akan menampilkan hasil berupa kesimpulan klasifikasi. Metode *Naïve Bayes* memiliki tingkat akurasi lebih tinggi dengan persentase nilai sebesar 87,5%. Sedangkan metode *K-Nearest Neighbor* nilai akurasi dengan nilai sebesar 71,25%

Daftar Pustaka

- [1] A. Sholikah, R. Rustiana, A. Y. Prodi, K. Masyarakat, and I. Artikel, “Faktor-Faktor yang Berhubungan dengan Status Gizi Balita di Pedesaan dan Perkotaan,” 2017. [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/phpj>
- [2] Kemenkes, “Hasil Studi Kasus Gizi Indonesia (SSGI)”, 2021. [Online]. Available: <http://www.badankebijakan.kemkes.go.id/buku-saku-hasil-studi-status-gizi-indonesia-ssgi-tahun-2021/>
- [3] M. Devi, “Analisis Faktor-Faktor Yang Berpengaruh Terhadap Status Gizi Balita Di Pedesaan,” 2010. [Online]. Available: <http://www.>
- [4] A. Jurnal *et al.*, “Perbandingan Metode K-Nearest Neighbors dan Naïve Bayes Classifier Pada Klasifikasi Status Gizi Balita di Puskesmas Muara Jawa Kota Samarinda,” *Adopsi Teknologi dan Sistem Informasi (ATASI)*, vol. 1, 2022, doi: 10.30872/atasi.v1i1.25.
- [5] B. Yulia, L. Fahik, B. S. Djahi, N. D. Rumlaklak, and J. I. Komputer, “Data Mining Untuk Klasifikasi Status Gizi Desa Di Kabupaten Malaka Menggunakan Metode K-Nearest Neighbor,” *J-ICON*, vol. 6, no. 1, pp. 1–7, 2018.
- [6] N. Epriliana, A. Putri, D. Syauqy, and M. H. Hanafi, “Sistem Klasifikasi Status Gizi Bayi dengan Metode K-Nearest Neighbor Berbasis Sistem Embedded,” 2017. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [7] M. Y. Titimeidara dan W. Hadikurniawati, “Monica Yoshe Titimeidara Implementasi Metode Naive Bayes Implementasi Metode Naive Bayes Classifier Untuk Klasifikasi Status Gizi Stunting Pada Balita.”
- [8] D. Simanjuntak dan A. Sindar, “Sistem Pakar Deteksi Gizi Buruk Balita Dengan Metode Naïve Bayes Classifier,” Online, 2019