

Bab II. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Tabel 2.1. Penelitian Terdahulu

Judul	Penulis Dan Tahun	Tujuan	Metode	Temuan	Kesamaan
Membandingkan algoritma <i>K-Nearest Neighbor</i> dan <i>Naïve Bayes</i> dalam menentukan status Gizi bayi dengan umur dibawah lima tahun di pusat Kesehatan masyarakat Muara Jawa di kota samarinda.	Jurnal A, Rizky Yuliansyah M, 2022	Mencari metode terbaik dari kode dua metode yang dibandingkan untuk menentukan status gizi balita	<i>K-Nearest Neighbor</i> dan <i>Naïve Bayes</i>	<i>Nearest Neighbor</i> lebih baik dari pada <i>Naïve Bayes</i> Setelah dilakukan proses pengukuran berdasarkan 3 indikator yaitu Akurasi, Presisi, <i>Recall</i> ,	Implementasi dua metode Yaitu <i>K-Nearest Neighbor</i> dan <i>Naïve Bayes</i>
Menentukan status gizi balita di daerah Kabupaten malaka menggunakan klasifikasi <i>Data Mining</i> dengan metode <i>K-Nearest Neighbor</i>	Yulia B, Fahik L, 2018	Menerapkan algoritma <i>K-Nearest Neighbor</i> apakah bisa memecahkan masalah pada penentuan status gizi balita	<i>K-Nearest Neighbor</i>	Algoritma <i>K-Nearest Neighbor</i> dapat melakukan proses klasifikasi status gizi balita presentase sistem sebesar 93.10%.	Metode <i>K-Nearest Neighbor</i>

Klasifikasi penentuan status gizi balita dengan algoritma <i>K-Nearest Neighbor</i> berdasarkan <i>Embedded System</i>	Epriliana N, Putri A, 2017	Membangun sebuah sistem <i>embedded</i> untuk mengklasifikasi status gizi balita	<i>K-Nearest Neighbor</i>	Sistem yang dibuat bisa bekerja dengan baik berdasarkan keakurasian sensor berat dengan nilai 97,23%.	<i>K-Nearest Neighbor</i>
Penerapan metode <i>Naïve Bayes</i> untuk mengklasifikasi status gizi balita	Titimeidara M, Hadikurniawati W, 2021	Menerapkan Metode <i>Naïve Bayes</i> apakah bisa memecahkan masalah pada penentuan status gizi balita	<i>Naïve Bayes</i>	Metode <i>Naïve Bayes</i> Mampu menentukan status gizi balita	<i>Naïve Bayes</i>
<i>Expert System</i> untuk mendeteksi gizi buruk pada balita menggunakan <i>Naïve Bayes</i>	Simanjuntak D, Sindar A, 2019	Membangun sebuah sistem pakar untuk mengklasifikasi status gizi balita dengan metode <i>Naïve Bayes</i>	<i>Naïve Bayes</i>	Metode <i>Naïve Bayes</i> mampu mendeteksi gizi buruk balita.	<i>Naïve Bayes</i>
Menentukan model terbaik dari metode <i>Naïve Bayes</i> untuk penentuan status gizi balita	Fitri A, Saptono R, Widya Sihwi S, 2015	Mengoptimalkan <i>Naïve Bayes Classifier</i> pada studi kasus penilaian status gizi dengan mempertimbangkan independensi parameter, dengan pemodelan data	<i>Naïve Bayes</i>	Kesimpulan penelitian ini membuktikan bahwa <i>Backward Feature Selection</i> bisa dimanfaatkan untuk membuat model menurut nilai korelasi yang digunakan dalam	<i>Naïve Bayes</i>

algoritma *Cosine*

Similarity.



Penelitian pertama membandingkan metode *K-Nearest Neighbor* dengan *Naïve 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 *Naïve Bayes* dengan skor if lebih besar (Jurnal et al., 2022)

Penelitian selanjutnya 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 $k = 7$ (Yulia et al., 2018)

Penelitian ketiga 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% (Epriliana et al., 2017)

Penelitian keempat menentukan status gizi balita dengan mengimplementasikan metode *Naïve Bayes*. Hasil dari penelitian ini menghasilkan tingkat akurasi dari 300 data dengan nilai sebesar 88% (Titimeidara & Hadikurniawati, n.d.,2021)

Penelitian kelima algoritma *Naïve Bayes* digunakan untuk membuat sistem pakar sebagai pendeteksi status gizi balita berdasarkan perhitungan antropometri. Hasil penelitian ini menunjukkan tingkat kinerja dengan hasil 0.0013168617 (Simanjuntak & Sindar, 2019)

Penelitian berikutnya menentukan model paling baik dari algoritma *Naïve Bayes* untuk penentuan status gizi balita. Kesimpulan penelitian ini membuktikan bahwa *Backward Feature Selection* bisa dimanfaatkan untuk membuat model menurut nilai korelasi yang digunakan dalam algoritma *Cosine Similarity*. Hasil dari tingkat akurasi sebesar 94,8% (1754-3912-1-PB, n.d.,2015)

2.2 Status Gizi

Status gizi merupakan kondisi tubuh yang terjadi akibat dampak mengonsumsi makanan dan zat tambahan yang mengandung zat gizi. Zat gizi sangat dibutuhkan tubuh untuk tumbuh kembang manusia secara maksimal. Status gizi dapat ditentukan melalui tes laboratorium dan tes antropometri, Pengukuran antropometri adalah pengukuran yang digunakan sebagai indikator penentuan status gizi anak sebagai pedoman. Status gizi yang masuk dalam klasifikasi buruk tidak hanya terbatas pada kurangnya kebutuhan gizi, kelebihan asupan sumber energi bisa membuat anak masuk dalam kategori obesitas. Factor-faktor yang menentukan status gizi balita seperti makanan, status social, pendapatan dan keadaan geografis (Riang Toby et al., 2021)

Pada sistem informasi yang didesain, akan menghasilkan dua klasifikasi yaitu klasifikasi NORMAL dan GIZI BURUK. Kategori status gizi dapat dilihat dalam table acuan berdasarkan standar antropometri.

Table 2.2 Kategori Status Gizi dan Ambang Batas (Kemenkes, 2020)

Indeks	Kategori Status Gizi	Ambang Batas
Berat Badan Balita Berdasarkan Umur (BB/U)	Kelebihan berat badan	$> +1SD$
	Berat Badan Normal	$-2 SD \text{ sd } + 1 SD$
	Berat Badan Kurang (Underweight)	$-3 SD \text{ sd } < -2 SD$
	Berat Badan Sangatlah Kurang (Severely Underweight)	$< -3SD$
	Tinggi Badan atau Panjang Balita Berdasarkan Umur (TB/U)	Tinggi
Tinggi Badan atau Panjang Balita Berdasarkan Umur (TB/U)	Normal	$-2 SD \text{ sd } + 1 SD$
	Pendek (Stunted)	$-3 SD \text{ sd } < -2 SD$
	Sangat Pendek (Severely Stunted)	$< -3SD$
Berat Badan Balita Berdasarkan Tinggi atau Panjang Badan Balita (BB/T Atau BB/P)	Obesitas (Obese)	$> +3SD$

Gizi Lebih (Overweight)	$> +2 SD \text{ sd} + 3 SD$
Resiko Gizi Berlebih (Possible Risk of Overweight)	$> +1 SD \text{ sd} + 2 SD$
Gizi Normal	$-2 SD \text{ sd} + 1SD$
Gizi Kurang (Wasted)	$-3 SD \text{ sd} < -2SD$
Gizi Buruk (Severely Wasted)	$< -3 SD$

2.3 Klasifikasi

Merupakan proses mencari pola atau model yang membedakan dan menggambarkan kelas pada data set. Klasifikasi digunakan untuk mengelompokkan data ke dalam kelas atau grup yang paling sesuai berdasarkan kriteria spesifik. Tujuan utama adalah prediksi kelas dari sebuah objek yang status labelnya belum ditentukan (Zhang et al., 2017)

2.4 K-Nearest Neighbor

Key Near Neighbor bekerja dengan mencari tetangga terdekat sebagai patokan untuk menentukan nilai dari data baru. *Offset* induktif yang dimiliki pada setiap label kelas dari nilai k dengan nilai label kelas yang akan di uji dan dicari kemiripannya dengan tetangga terdekat. *K-Nearest Neighbor* juga disebut *lazy* algoritma karena tidak menggunakan *training sample* untuk kebutuhan belajar (*generalisasi*) atau hanya sedikit untuk belajar.

Sebuah algoritma yang mencari sebuah titik pada kisi lalu menentukan sebuah titik yang terletak di grup A atau grup B, setelah itu melihat titik apa saja yang berada di dekatnya. Jarak yang dipilih ditentukan secara acak, hal yang paling penting adalah mengambil data sampel. Hitung mayoritas titik di setiap grup, jika ditemukan mayoritas berada titik grup A maka dapat disimpulkan klasifikasi titik data berada di A dan sebaliknya apabila mayoritas titik berada di grup B maka klasifikasi berada di grup B (Xing & Bei, 2020)

Eulidean distance merupakan rumus utama dari metode *K-Nearest Neighbor* agar dapat mengklasifikasi data.

$$d(p_i, q_i) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (1)$$

Keterangan :

$d(p_i, q_i)$: Nilai jarak antara p_i dan q_i

p_i : Nilai data uji ke i pada variabel pertama.

q_i : Nilai data latih ke i pada variabel pertama.

2.5 Naïve Bayes

Naïve Bayes merupakan salah satu metode klasifikasi yang efektif dan efisien. *Naïve Bayes* termasuk algoritma pembelajaran induktif yang digunakan dalam *data mining* dan *machine learning*. Hasil didapatkan dari menghitung probabilitas untuk mendapatkan asumsi akurat. Estimasi dari probabilitas merupakan hal yang krusial dari algoritma ini. Ada tiga jenis pemodelan *Naïve Bayes* yaitu *Multinomial Naïve*, *Gaussian Naïve*, dan *Bernoulli Naïve*. *Multinomial Naïve* digunakan saat kondisi fitur-fitur didapatkan dari distribusi *simple Multinomial*, model ini paling cocok digunakan pada data diskrit. *Gaussian Naïve* merupakan pemodelan paling sederhana dari *Naïve Bayes* cara kerja pemodelan ini dengan asumsi dari setiap label data yang diambil dari *Gaussian Naïve*. Pemodelan terakhir adalah *Bernoulli Naïve*, model ini digunakan saat fitur diasumsikan berbentuk (*true* atau *false*) atau sering disebut *biner 0* atau *1* (H. Zhang, n.d.)

Perhitungan *Naïve Bayes* menggunakan rumus Probabilitas untuk variabel diskrit yaitu

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \quad (2)$$

Keterangan :

B : Data dengan label atau *kelas* yang akan dicari.

A : Hipotesis atau asumsi sementara dari kelas data.

$P(A|B)$: Probabilitas hipotesis A berdasarkan Keadaan B

$P(A)$: Probabilitas hipotesis A

$P(A|B)$: Probabilitas hipotesis B berdasarakan keadaan A

$P(B)$: Probabilitas hipotesis B

Perhitungan *Naïve Bayes* menggunakan rumus Probabilitas untuk variabel kontinyu yaitu :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3)$$

Keterangan :

$f(x)$: Fungsi x.

π : Nilai Pi.

σ : Nilai Standar Deviasiasi.

μ : Nilai Rata-Rata.

e : Nilai Eksponen.

