

IMPLEMENTASI ALGORITMA *ITERATIVE DICHOTOMIZER 3* UNTUK DIAGNOSIS PENYAKIT GASTROENTERITIS ANAK BERBASIS *WEB*

M. Karman Habibi¹⁾, Veri Julianto²⁾, M. Najamudin Ridha³⁾, Fathurrahmani⁴⁾

¹Program Studi Teknologi Informasi, Politeknik Negeri Tanah Laut

^{1,2,3,4}Komputer Dan Bisnis, Politeknik Negeri Tanah Laut

E-mail: veri@politala.ac.id

Abstract

Gastroenteritis is a disease that affects the digestive system, causing symptoms such as nausea, vomiting, diarrhea, abdominal cramps, and dehydration. Lack of knowledge and delayed diagnosis often worsen the condition of children affected by gastroenteritis. Therefore, early awareness is needed to prevent wider spread. The ID3 algorithm, which performs an exhaustive search of all possible decision trees, is used in this study to predict gastroenteritis in children. Data from 69 affected children and 61 unaffected children at RSUD Hadji Boejasin Pelaihari shows that the ID3 algorithm achieved an accuracy of 86.6%, precision of 92.3%, and sensitivity of 80%. The implementation of this algorithm in a web-based application is expected to provide useful information for the early diagnosis of gastroenteritis.

Keywords: *Gastroenteritis, Diagnosis, Iterative Dichotomizer 3, Decision Tree, Confusion Matrix*

PENDAHULUAN

Gastroenteritis merupakan penyakit menular yang menyerang sistem pencernaan tubuh, terutama usus kecil dan lambung, melalui virus dan bakteri. Biasanya, virus atau bakteri yang menyerang bernama *norovirus*, *rotavirus*, dan *campylobacter*. Gastroenteritis memunculkan gejala berupa peradangan pada lambung (*gaster*) dan usus kecil (*entero*). Peradangan ini menyebabkan gejala berupa mual, muntah, diare, kram perut, dan dehidrasi yang berlebihan (Dewa Made Krisnayana et al., 2020).

Berdasarkan data dari Riset Kesehatan Dasar Indonesia, prevalensi tertinggi penyakit gastroenteritis tercatat di lima provinsi, yaitu Aceh (10,2%), Papua (9,6%), DKI Jakarta (8,9%), Sulawesi Selatan (8,1%), dan Banten (8,0%) (Nari Jois, 2019). Tingkat prevalensi yang tinggi disebabkan oleh kurangnya pemahaman masyarakat mengenai gejala penyakit gastroenteritis. Hal ini terutama terjadi di kalangan orang tua, yang sering kali menyebabkan kesalahan dalam mendiagnosa dan penanganan. Kesalahan tersebut dapat mengakibatkan anak mengalami dehidrasi secara berlebihan.

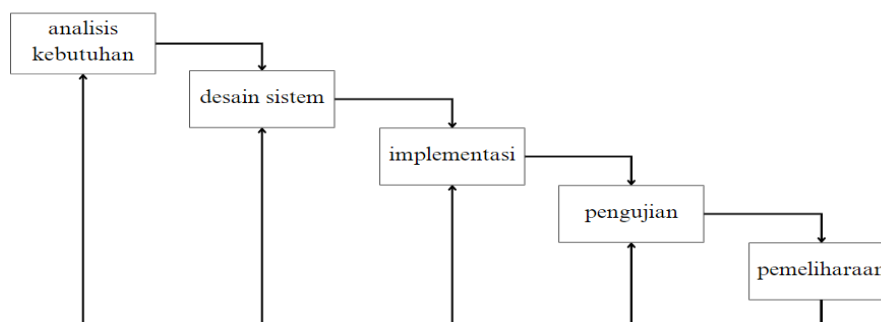
Oleh sebab itu diperlukan sistem yang dapat memprediksi penyakit gastroenteritis serta penanganan akan penyakit tersebut.

Data mining ada proses menemukan hubungan dan pola yang signifikan dengan melakukan analisis dari sebuah data (Nabila et al., 2021). Salah satunya dengan algoritma *Iterative Dichotomizer 3* (ID3) adalah algoritma pembelajaran pohon keputusan yang paling dasar. Algoritma ini melakukan pencarian *greedy* atas semua pohon keputusan yang mungkin. Algoritma ID3 dapat diimplementasikan menggunakan fungsi rekursif (Anggriawan, 2021). Algoritma ID3 dapat digunakan untuk mengklasifikasikan seseorang terkena suatu penyakit, berikut penelitian-penelitian terdahulu yang relevan antara lain: Implementasi ID3 untuk penyakit campak (Srimenganti et al., 2018), sistem pakar untuk memprediksi penyakit sapi menggunakan pohon keputusan ID3 (Anggriawan, 2021), Penerapan algoritma *iterative dichotomizer 3* (ID3) dalam mendiagnosa Kesehatan kehamilan (Rahmawati et al., 2019).

Pada penelitian ini akan menerapkan metode ID3 dalam memprediksi penyakit gastroenteritis berdasarkan data-data yang di dapat dari Rumah Sakit. Metode ini akan memberikan gambaran awal kepada pengguna sistem sehingga mampu mendidagnosa secara mandiri sebelum melakukan pengobatan ke dokter. Sistem ini memberikan referensi dalam pengambil keputusan terhadap gejala-gejala awal yang dialami oleh pengguna.

METODE PENELITIAN

Metode pengembangan sistem pada penelitian ini menggunakan metode waterfall. Tahapan dalam pengembangan model waterfall yaitu analisis kebutuhan, desain sistem, implementasi, pengujian dan pemeliharaan dapat dilihat pada Gambar 1.



Gambar 1. Model Pengembangan Sistem Waterfall

Analisis Kebutuhan

Analisis kebutuhan bertujuan untuk memahami kebutuhan pengguna dan menentukan spesifikasi perangkat lunak yang tepat (Badrul, 2021). Dalam membangun aplikasi berbasis web untuk diagnosis gastroenteritis pada anak, langkah pertama adalah menganalisis data yang diperlukan agar sistem berfungsi dengan baik. Penulis mempelajari literatur terkait untuk mengidentifikasi masalah dalam memprediksi gastroenteritis.

Desain Sistem

Pada tahap ini, pengembang merancang sistem untuk membantu menentukan persyaratan perangkat keras (*hardware*) dan sistem, serta mendefinisikan arsitektur sistem secara keseluruhan (Abdul Wahid Aceng, 2020). Pada tahap ini meliputi perancangan basis data dan perancangan sistem. Pada tahap perancangan basis data menggunakan ERD dan untuk perancangan sistem menggunakan UML.

Implementasi

Pengkodean adalah proses mengubah desain menjadi bahasa yang dipahami oleh komputer, mengimplementasikan desain menjadi sistem yang nyata, termasuk basis data, perhitungan ID3, tampilan antarmuka, dan fitur sistem. Penulis menggunakan bahasa pemrograman PHP dengan database MySQL pada tahap ini.

Pengujian

Serangkaian pengujian dilakukan untuk memastikan bahwa setiap komponen berfungsi sesuai harapan yang telah ditetapkan sebelumnya. Terakhir yaitu melakukan pengujian black-box untuk fungsionalitas dan confusion matrix untuk menguji akurasi, presisi dan *recall*.

Pemeliharaan

Pemeliharaan sistem mencakup aktivitas penting seperti perbaikan kesalahan, yaitu mengidentifikasi dan memperbaiki bug adaptasi sistem untuk menyesuaikan dengan perubahan teknologi.

HASIL DAN PEMBAHASAN

Analisis Sistem Yang Diusulkan



Gambar 2. Analisis sistem yang diusulkan

Publik atau orang tua anak mengakses sistem melakukan konsultasi online dan mendapatkan hasil konsultasi berupa diagnosa apakah anak terkena gastroenteritis atau tidak beserta penanganannya. Admin mengatur keseluruhan sistem yaitu berupa data gejala, data pasien, data sampel, perhitungan klasifikasi ID3 dan yang lainnya.

Perhitungan Iterative Dichotomizer 3

Perhitungan ID3 menggunakan data asli yang didapat dari rekam medis pasien anak yang terkena gastroenteritis dan pasien anak yang tidak terkena penyakit gastroenteritis. Adapun data awal didapatkan dari RSUD Hadji Boejasin Pelaihari yaitu data rekam medis sebanyak 130 data yang dibagi untuk data training berjumlah 100 dan data testing berjumlah 30, data awal dapat dilihat pada Tabel 1.

Tabel 1
Data awal

Nama	Gejala	Penyakit
Pasien 1	Demam, Mual, Muntah, Nyeri Perut	Dengue Fever
Pasien 2	Lemas, Mual, Demam, Muntah	Hipokalemia
Pasien 3	Demam, Mual, Muntah, Batuk	Obs Febris
Pasien 4	Demam, Muntah, Bab Cair	Gastroenteritis
Pasien 5	Batuk, Demam, Mual, Muntah	Gastroenteritis

Perhitungan entropi total dilakukan dengan menghitung jumlah kasus positif dan negatif gastroenteritis, lalu menghitung entropi setiap atribut dan gain untuk menemukan gain terbesar sebagai akar pohon.

$$Entropy(S) = \sum_{j=1}^k -p_i \log_2 p_i \quad (1)$$

dengan keterangan sebagai berikut,

S = himpunan kelas klasifikasi.

k = banyaknya kelas klasifikasi.

Pi = proporsi untuk kelas i.

$Entropy(\text{Total}) = (-54/100 * \log_2(54/100)) + (-(46/100) * \log_2(46/100)) = 0,99537843$.

Selanjutnya menentukan gain (demam) rumus mencari gain adalah sebagai berikut:

$$Gain(A) = Entropy(S) - \sum_{i=1}^k \frac{|S_i|}{|S|} \times Entropy(S_i) \quad (2)$$

dengan keterangan sebagai berikut, A = atribut.

S = ruang (data) sampel digunakan untuk training.

A = atribut.

$|S_i|$ = jumlah sampel untuk nilai V.

$|S|$ = jumlah seluruh sampel data.

$Entropy(S_v)$ = entropy untuk sampel-sampel yang memiliki nilai v.

Mengitung gain untuk demam.

$$\begin{aligned} \text{Gain (Demam)} &= 0,99537843 - \left(\left(\frac{84}{100} \times 0,999591034 \right) + \left(\frac{16}{100} \times 0,69621226 \right) \right) \\ &= 0,044328008. \end{aligned}$$

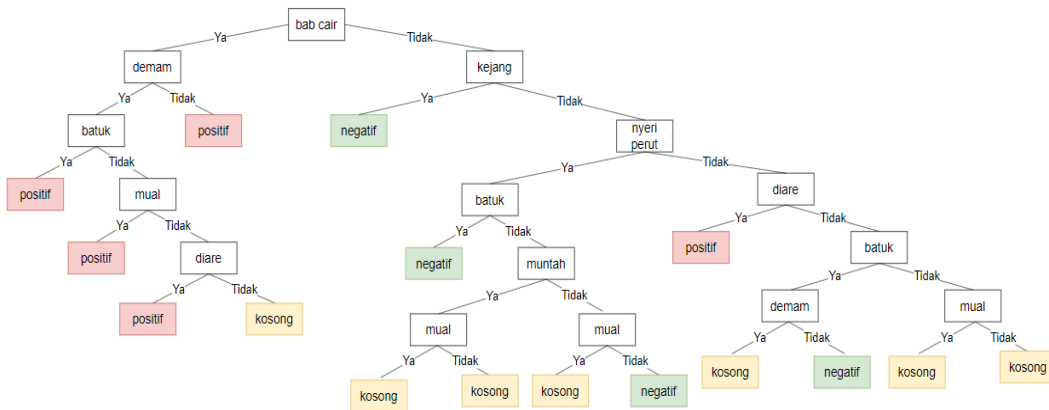
Untuk penentuan node dilakukan perhitungan seluruh gain dari setiap atribut dapat dilihat pada Tabel 2.

Tabel 2
Perhitungan Gain

Gejala	Ket	Jumlah	Positif	Negatif	Entropy	Gain
Total		100	54	46	0,995378439	
Demam	Ya	84	41	43	0,999591034	0,044328008
	Tidak	16	13	3	0,69621226	
Muntah	Ya	71	43	28	0,967559482	0,030720696
	Tidak	29	11	18	0,957553484	
Mual	Ya	34	13	21	0,959686894	0,037345145
	Tidak	66	41	25	0,957181439	
Batuk	Ya	33	14	19	0,98337619	0,019175402
	Tidak	67	40	27	0,972669991	
Pilek	Ya	2	0	2	0	0,022751955
	Tidak	98	54	44	0,992476004	
Nyeri	Ya	31	7	24	0,770629069	0,133333269
Perut	Tidak	69	47	22	0,903116172	
Bab	Ya	43	42	1	0,159350063	0,503639997
Cair	Tidak	57	12	45	0,74248757	0,095553146
	Ya	8	0	8	0	
Kejang	Tidak	92	54	38	0,978070971	0,041058565
	Ya	8	7	1	0,543564443	
Diare	Tidak	92	47	47	0,990037737	0,012346854
	Ya	9	3	6	0,918295834	
Nafas	Tidak	91	51	40	0,989434022	

Pohon Keputusan

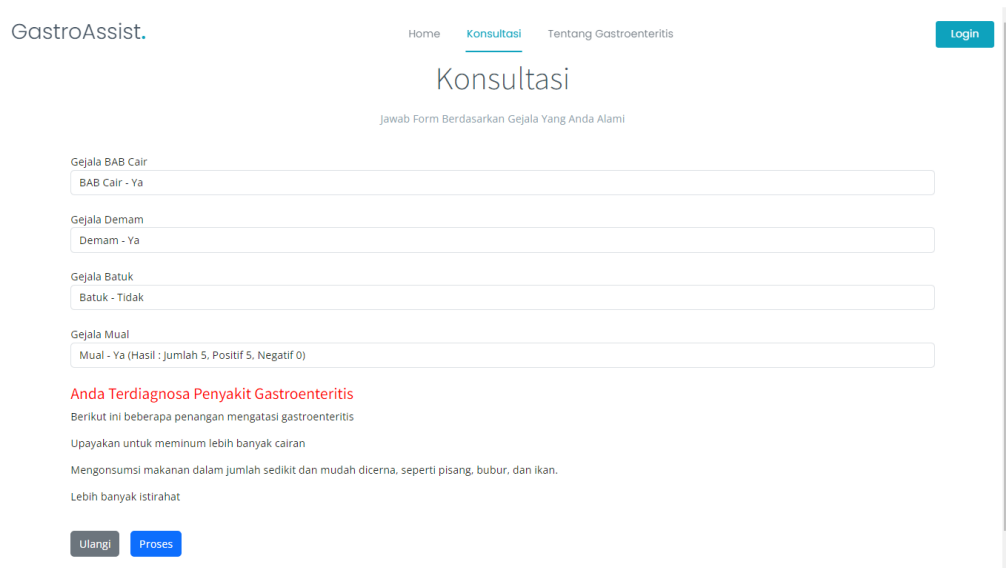
Pembuatan pohon keputusan didapatkan dari nilai gain tertinggi yang dijadikan root pada pohon sehingga dapat menghasilkan pohon keputusan yang memiliki rule akhir hasil pohon keputusan dapat dilihat pada Gambar 3.



Gambar 3. Pohon Keputusan

Implementasi Sistem

Halaman landing page konsultasi adalah halaman yang dapat digunakan oleh pengguna umum untuk melakukan diagnosis awal penyakit gastroenteritis. Halaman ini berisi aturan dinamis yang diperoleh dari perhitungan ID3 pada sistem dapat dilihat pada Gambar 4.



Gambar 4. Landing Page Konsultasi

Pengujian Confusion Matrix

Confusion Matrix adalah metode untuk mengukur performa dalam data mining, membandingkan hasil klasifikasi prediksi dengan hasil klasifikasi aktual. (Dewa Made Krisnayana et al., 2020). Pengujian *Confusion Matrix* dilakukan untuk mengevaluasi kinerja model klasifikasi yang dibangun. *Confusion matrix* memberikan gambaran yang jelas tentang kemampuan model dalam mengklasifikasikan data dengan benar dan salah atau diagnosis positif dan negatif untuk penyakit gastroenteritis. Digunakan 30 data *testing* yang memiliki label positif 15 dan negatif 15.

Hasil prediksi sistem dari model ID3 yang sudah dibuat, setelah dilakukan uji menggunakan 30 data *testing*, menunjukkan hasil sebagai berikut: 12 true positive, 14 true negative, 1 false positive, dan 3 false negative dapat dilihat pada Tabel 3.

Tabel 3
Total Hasil Prediksi

	Prediksi Positif	Prediksi Negative
Kebenaran Positif	TP: 12	FN: 3
Kebenaran Negative	FP: 1	TN: 14

Selanjutnya dilakukan perhitungan confusion matrix pada tiap-tiap performansi. Nilai akurasi sebesar 86,6% menunjukkan kemampuan model dalam mengklasifikasikan data dengan benar. Nilai presisi sebesar 92,3% menunjukkan bahwa sebagian besar prediksi positif model adalah benar. Nilai recall sebesar 80% menunjukkan kemampuan model dalam mengenali data positif sebenarnya, dapat dilihat pada Tabel 4.

Tabel 4
Hasil pengujian dengan confusion matrix

Confusion Matrix	Rumus	Perhitungan	Hasil
Akurasi	$\frac{TP + TN}{TP + TN + FP + FN}$	$\frac{12 + 14}{12 + 14 + 1 + 3}$	86,6%
Presisi	$\frac{TP}{TP + FP}$	$\frac{12}{12 + 1}$	92,3%
Recall	$\frac{TP}{TP + FN}$	$\frac{12}{12 + 3}$	80%

SIMPULAN

Metode Iterative Dichotomizer 3 berhasil diterapkan untuk memprediksi apakah seseorang terkena gastroenteritis atau tidak, menggunakan gejala pasien anak yang

terjangkit dan tidak terjangkit gastroenteritis. Hasil performansi metode ID3 menggunakan Confusion Matrix dengan data training sebanyak 100 dan data testing sebanyak 30, menunjukkan akurasi sebesar 86,6%, presisi 92,3%, dan recall 80%, yang berarti aplikasi cukup akurat, andal, dan sensitif dalam diagnosis penyakit gastroenteritis.

DAFTAR PUSTAKA

- Abdul Wahid Aceng. (2020). *Analisis Metode Waterfall Untuk Pengembangan Sistem Informasi*.
- Anggriawan, D. (2021). Sistem Pakar Untuk Memprediksi Penyakit Pada Hewan Ternak Sapi Menggunakan Pohon Keputusan Id3. In *Jecsit* (Vol. 1, Issue 1).
- Badrul, M. (2021). *Penerapan Metode Waterfall Untuk Perancangan Sistem Informasi Inventory Pada Toko Keramik Bintang Terang*. 8(2).
- Dewa Made Krisnayana, I., Arya Mertasana, P., & Sudarma, M. (2020). Aplikasi Sistem Pakar Diagnosa Penyakit Gastroenteritis Berbasis Android Dengan Metode Classification And Regression Tree. In *Putu Arya Mertasana* (Vol. 7, Issue 3).
- Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (Jtsi)*, 2(2), 100.
- Nari Jois. (2019). 366-762-1-Pb. *Global Health Science*, 4(3), 159–164.
- Rahmawati, A., Wintana, D., Ayu Ambarsari, D., Nusa Mandiri Jakarta Jl Kramat Raya No, S., & Kwitang Kec Senen, K. (2019). *Penerapan Algoritma Iterative Dichotomiser Three (Id3) Dalam Mendiagnosa Kesehatan Kehamilan*.
- Srimenganti, I., Taufik, I., Mulyana, E., Informatika, T., Gunung, S., Bandung, D., & Elektro, T. (2018). *Implementasi Algoritma Decision Tree (Id3) Untuk Penyakit Campak*.