

**Penerapan Algoritma Naive Bayes Untuk Prediksi Penyakit Diare Pada Balita****Muhammad Akbar Wahyudi<sup>1</sup>, Ahmad Naufal Jiddan<sup>2</sup>, Ramdhan Saepul Rohman<sup>3</sup>, Eva Marsusanti<sup>4</sup>**Universitas Bina Sarana Informatika<sup>1,2,3,4</sup>makbarwhyd@gmail.com<sup>1</sup>, naufaljiddan7@gmail.com<sup>2</sup>, ramdhan.rpe@bsi.ac.id<sup>3</sup>, eva.emr@bsi.ac.id<sup>4</sup>

Diterima (30-08-2023)	Direvisi (21-09-2023)	Disetujui (20-10-2023)
--------------------------	--------------------------	---------------------------

**Abstrak** - Penyakit diare adalah kondisi yang sering terjadi dimana seseorang mengalami frekuensi buang air besar yang tinggi dengan tinja yang cair. Penyakit ini bisa disebabkan oleh berbagai faktor, termasuk infeksi virus, bakteri, atau parasit, konsumsi makanan atau air yang terkontaminasi, atau efek samping dari obat-obatan tertentu. Pada tahun 2020, diare masih menjadi masalah utama yang menyebabkan 14,5% kematian. Pada kelompok anak balita (12 – 59 balita), kematian akibat diare sebesar 4,55%. Sudah banyak peneliti yang melakukan penelitian untuk memprediksi penyakit diare pada balita. Penelitian ini dilakukan di Klinik Gema Medical Center Cabang Cicurug terhadap pasien diare dengan tujuan menghasilkan perhitungan terbaik berupa akurasi tentang diagnosis penyakit diare yang nantinya dapat membantu dalam prediksi diagnosis lebih dini. Adapun metode yang digunakan yaitu metode Naive Bayes untuk meningkatkan akurasi prediksi yang diaplikasikan terhadap data pasien yang dinyatakan Ya diare dan Tidak diare. Setelah dilakukan pengujian dengan metode naive bayes dengan Attribute yang digunakan yaitu jenis kelamin usia (bulan) frekuensi bab, konsistensi tinja, mata, keinginan untuk minum, dan keadaan umum. Sehingga hasil yang didapat dari confusion matrix dan kurva ROC dengan menggunakan Naive Bayes diperoleh accuracy sebesar 98,57% dengan nilai AUC 1,000 dengan tingkat diagnosa excellent classification karena hasil AUC-nya antara 0,90-1,00.

**Kata Kunci** : Penyakit Diare, Naive Bayes, Prediksi, Data Mining

**Abstract** - Diarrhea is a condition that often occurs where a person experiences a high frequency of defecation with loose stools. This disease can be caused by various factors, including viral, bacterial or parasitic infections, consumption of contaminated food or water, or side effects of certain medications. In 2020, diarrhea is still a major problem causing 14.5% of deaths. In the group of children under five (12 – 59 toddlers), deaths due to diarrhea were 4.55%. Many researchers have conducted research to predict diarrhea in toddlers. This research was carried out at the Gema Medical Center Clinic, Cicurug Branch, on diarrhea patients with the aim of producing the best calculations in the form of accuracy regarding the diagnosis of diarrheal diseases which can later help in predicting earlier diagnoses. The method used is the Naive Bayes method to increase the accuracy of predictions applied to patient data that states Yes and No diarrhea. After testing using the naive Bayes method, the attributes used were gender, age (months), frequency of bowel movements, stool consistency, eyes, desire to drink, and general condition. So the results obtained from the confusion matrix and ROC curve using Naive Bayes obtained an accuracy of 98.57% with an AUC value of 1,000 with an excellent classification diagnostic level because the AUC results were between 0.90-1.00.

**Keywords:** Diarrheal, Naive Bayes, Prediction, Data Mining

**I. PENDAHULUAN**

Problema penyakit merupakan salah satu faktor utama dalam usaha perbaikan kesehatan masyarakat. Hal tersebut disebabkan jumlah pasien yang terus bertambah bahkan berdampak pada kematian jika tidak segera ditangani sehingga diperlukan persiapan dalam menghadapi situasi peningkatan jumlah kasus. Salah satu penyakit endemis yang sering terjadi di Indonesia adalah diare.

Penyakit diare adalah kondisi yang sering terjadi dimana seseorang mengalami frekuensi

buang air besar yang tinggi dengan tinja yang cair (Sinum, 2021). Penyakit ini bisa disebabkan oleh berbagai faktor, termasuk infeksi virus, bakteri, atau parasit, konsumsi makanan atau air yang terkontaminasi, atau efek samping dari obat-obatan tertentu (Hartati & Nurazila, 2018).

Kesehatan merupakan bagian kehidupan manusia yang sangat penting, sebab setiap manusia dapat terganggu kesehatannya. Tidak terkecuali anak yang mudah terkena penyakit, sehingga sangat penting bagi orang tua untuk

mendapatkan segala bentuk informasi tentang penyakit anak yang sedang diderita. Kebanyakan orang tua pada umumnya sering tidak mengetahui penyakit yang gejala-gejalanya timbul pada tubuh anak. Informasi mengenai gejala-gejala penyakit pada tubuh anak sangat penting untuk diketahui orang tua agar orang tua mampu melakukan tindakan awal yang tepat. Hal ini dikarenakan jika terjadi kesalahan ataupun keterlambatan dalam mengenali gejala-gejala serta jenis penyakit pada anak, dapat menyebabkan kesalahan ataupun keterlambatan dalam penanganan pengobatannya hingga dapat menyebabkan kematian (Bahri, Marisa Midyanti, & Hidayati, 2018).

Salah satu penyakit yang sering terjadi pada anak yaitu diare. Menurut Kemenkes pada tahun 2020, diare masih menjadi masalah utama yang menyebabkan 14,5% kematian. Pada kelompok anak balita (12 – 59 balita), kematian akibat diare sebesar 4,55% (Kesehatan, 2023). Sedangkan berdasarkan penyakit menular, diare merupakan penyebab kematian peringkat ke-3 setelah TB dan Pneumonia (Rahman & Muzakir, 2020).

Prediksi penyakit diare menjadi penting dalam bidang kesehatan karena memungkinkan langkah-langkah pencegahan yang diperlukan untuk dikembangkan dan diimplementasikan. Salah satu pendekatan yang dapat digunakan dalam prediksi penyakit diare adalah algoritma Naive Bayes.

Algoritma Naive Bayes adalah salah satu metode pada teknik klasifikasi untuk mengatasi ketidakpastian data dan termasuk dalam classifier statistik yang dapat memprediksi probabilitas keanggotaan class (Syarah, Wati, & Puspitasari, 2022). Metode ini didasarkan pada asumsi bahwa semua fitur yang ada dalam dataset adalah independen satu sama lain, yang sering disebut sebagai asumsi "naive" (Laila Sari, Saputra, & Gemasih, 2022). Meskipun asumsi ini jarang terpenuhi dalam dunia nyata, algoritma Naive Bayes tetap memberikan hasil yang baik dalam banyak kasus (Loelianto, Thayf, & Angriani, 2020).

Berdasarkan permasalahan diatas dapat disimpulkan bahwa algoritma Naive Bayes adalah metode yang berguna untuk memprediksi penyakit diare berdasarkan gejala-gejala yang dialami oleh pasien. Dengan mengumpulkan dan menganalisis data gejala serta mengaplikasikan rumus probabilitas Bayes, algoritma ini dapat membantu dalam memberikan prediksi yang akurat, yang dapat mendukung upaya pencegahan dan pengendalian penyakit diare.

**II. METODOLOGI PENELITIAN**

Jenis penelitian ini menggunakan pendekatan analisa kuantitatif Data penyakit diare balita menggunakan algoritma Naive Bayes. Naive

bayes merupakan metode yang tidak memiliki aturan, Naive Bayes menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi, dengan cara melihat frekuensi tiap klasifikasi pada data training. Klasifikasi Naive Bayes adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Klasifikasi bayesian didasarkan pada teorema bayes yang memiliki kemampuan klasifikasi serupa dengan decision tree dan neural network (Mutiara, 2020). Naive Bayes dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang dimasa depan dengan berdasarkan pengalaman dimasa sebelumnya. Metode Naive Bayes juga dinilai berpotensi baik dalam mengklasifikasikan dokumen dibandingkan dengan metode pengklasifikasian lain dalam hal akurasi (Wahyuningsih & Patima, 2018). Bayes rule digunakan untuk menghitung probabilitas suatu class. Algoritma Naive Bayes memberikan suatu cara mengkombinasikan peluang terdahulu dengan syarat kemungkinan menjadi sebuah formula yang dapat digunakan untuk menghitung peluang dari tiap kemungkinan yang terjadi. Bentuk umum dari teorema bayes seperti dibawah ini:

$$P(H | X) = (P(X | H) P(H)) / (P(X)) \dots\dots\dots(1)$$

Dimana:

- X : Data dengan class yang belum diketahui
- H : Hipotesis data X merupakan suatu class spesifik
- P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posterior probability)
- P(H) : Probabilitas hipotesis H (prior probability)
- P(X|H) : Probabilitas X berdasar kondisi pada hipotesis H
- P(X) : Probabilitas dari X

Naive bayes adalah penyederhanaan metode bayes. Teorema bayes disederhanakan menjadi:

$$P(H|X) = P(X|H) P(X) \dots\dots\dots(2)$$

Hasil dari masing-masing pengujian tersebut kemudian divalidasi menggunakan tools yang sama yaitu Rapidminer Studio 8.2 dengan memperoleh nilai akurasi. Sedangkan untuk evaluasi menggunakan dua measurment yaitu berdasarkan confusion matrix (CM) dan evaluasi kurva ROC.

Confusion matrix adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah

(Normawati & Prayogi, 2021). Confusion Matrix digunakan untuk menentukan persentase accuracy, precision dan recall. Pada pengukuran kinerja menggunakan Confusion Matrix, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Dimana Nilai True Negative (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan False Positive (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, True Positive (TP) merupakan data positif yang terdeteksi benar. False Negative (FN) merupakan kebalikan dari True Positive, sehingga data positif, namun terdeteksi sebagai data negative (Fauziningrum & Suryaningsih, 2021).

Sedangkan Kurva ROC adalah tool dua dimensi di dunia data mining yang digunakan untuk menilai kinerja klasifikasi yang menggunakan dua class keputusan, tiap objek dipetakan ke salah satu elemen dari himpunan pasangan, positif atau negatif. Pada gambar kurva ROC, True Positif rate diplot pada sumbu Y dan False Positif rate diplot pada sumbu X (Ridwan, 2020). Menurut Gorunescu Dalam metode klasifikasi data mining, hasil AUC dapat dibagi menjadi beberapa kelompok (Giovani, Ardiansyah, Haryanti, Kurniawati, & Gata, 2020):

1. 0,90-1,00 = Klasifikasi Sangat Baik
2. 0,80-0,90 = Klasifikasi Baik
3. 0,70-0,80 = Klasifikasi Sedang
4. 0,60-0,70 = Klasifikasi Buruk
5. 0,50-0,60 = Kegagalan

Penelitian ini dilakukan di Klinik Gema Medical Center Cabang Cicurug dengan menggunakan populasi data penyakit diare pada balita Tahun 2020-2023 sebanyak 181 data direduksi dengan menghilangkan duplikasi menjadi 68 data yang terdiri dari pasien diare balita sebanyak 37 data dan pasien yang tidak diare sebanyak 31 data.

Sedangkan untuk sampel pada penelitian ini adalah data rekap medik penyakit diare pada balita dengan hasil Ya dan Tidak, data tersebut bersifat intern yang belum pernah dipublikasikan oleh pihak Klinik. Dapat dilihat pada tabel berikut:

Tabel 1. Sampel Dataset

Ya	Tidak	Total Sampel
37	31	68

Sumber : (Klinik Gema Medical Center Cabang Cicurug, 2023)

Berikut Atribut dan Nilai Kategori Penyakit Diare Pada Balita, seperti terlihat pada tabel. 2 di bawah ini:

Tabel 2. Atribut dan Nilai Kategori Penyakit Diare Pada Balita

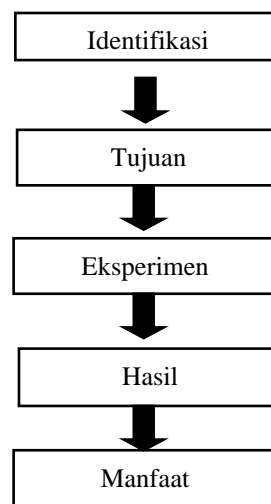
Atribut	Kategori
---------	----------

Jenis Kelamin	Perempuan Laki-Laki
Usia (Bulan)	< 6 6 – 11 12 – 23 24 – 35 36 – 47 48 – 59
Frekuensi	> 3 <= 3
Konsistensi Tinja	Cair Lembek Padat
Mata	Cekung Tidak Cekung
Keinginan Minum	Untuk Normal Ingin Minum Terus Malas Minum
Keadaan Umum	Gelisah Sadar Tidak Sadar
Hasil	Ya Tidak

Sumber : (Klinik Gema Medical Center Cabang Cicurug, 2023)

Untuk memprediksi penyakit diare pada balita maka peneliti menggunakan metode klasifikasi Naïve Bayes pada data rekam medik penyakit diare pada balita di Klinik Gema Medical Center Cabang Cicurug .

Kerangka penelitian dapat dilihat pada gambar berikut:



Sumber : (Putri, 2017)

Gambar 1. Kerangka Penelitian

Penjelasan Kerangka Penelitian:

1. Permasalahan  
Sering terjadi kesalahan ataupun keterlambatan dalam mengenali gejala-gejala serta jenis penyakit pada anak, dapat menyebabkan kesalahan ataupun keterlambatan dalam penanganan pengobatannya hingga dapat menyebabkan kematian.

2. Tujuan  
Tujuan penelitian ini adalah menghasilkan perhitungan terbaik yaitu berupa akurasi tentang diagnosis penyakit diare dengan menggunakan Algoritma Naive Bayes yang nantinya dapat membantu dalam prediksi diagnosis lebih dini.
3. Eksperimen  
Inputan data merupakan data Penyakit Diare Pada Balita di Klinik Gema Medical Center Cabang Cicurug, metode algoritma yang digunakan menggunakan Naive Bayes, dan implementasi/tools yang digunakan menggunakan Rapid Miner
4. Hasil  
Dapat memprediksi penyakit diare pada balita dengan menggunakan metode naive bayes.
5. Manfaat  
Dapat membantu pihak klinik dalam memprediksi penyakit diare pada balita.

**III. HASIL DAN PEMBAHASAN**

Dari data rekam medik yang telah didapat maka diambil sampel sebanyak 68 data yang terdiri dari pasien diare balita sebanyak 37 data dan pasien yang tidak diare sebanyak 31 data. sample data pasien, seperti terlihat pada tabel 3 halaman terakhir.

1. Eksperimen dan Pengujian Metode Naive Bayes

Pembuatan model naive bayes dilakukan pada dataset yang terdiri dari 8 atribut yang merupakan atribut dari diagnosa penyakit diare pada balita dan class yang merupakan hasil akhir prediksi.

Dalam membuat model naive bayes terlebih dahulu kita mencari nilai probabilitas hipotesis untuk masing-masing kelas P(H). Hipotesis yang ada yaitu pasien ya diare dan pasien tidak diare. Data training yang digunakan dengan total 68 data dengan 37 data pasien ya diare dan 31 data pasien tidak diare, perhitungan probabilitas seperti dibawah ini:

$P(\text{Ya Diare}) = 37:68 = 0,544118$

$P(\text{Tidak Diare}) = 31:68 = 0,455882$

Setelah probabilitas untuk setiap hipotesis diketahui langkah selanjutnya adalah menghitung probabilitas kondisi tertentu (probabilitas X) berdasarkan probabilitas tiap hipotesis (probabilitas H) atau dinamakan probabilitas prior dengan menggunakan naive bayes.

$P(\text{Jenis Kelamin Laki-Laki} | \text{Ya Diare}) = 16/37 = 0,432432$

Jadi, peluang Jenis Kelamin Laki-Laki Ya Diare sebesar 0,432432

$P(\text{Jenis Kelamin Laki-Laki} | \text{Tidak Diare}) = 16/31$

$= 0,516129$

Jadi, peluang Jenis Kelamin Laki-Laki Tidak Diare sebesar 0,516129

$P(\text{Jenis Kelamin Perempuan} | \text{Ya Diare}) = 21/37 = 0,567568$

Jadi, peluang Jenis Kelamin Perempuan Ya Diare sebesar 0,567568

$P(\text{Jenis Kelamin Perempuan} | \text{Tidak Diare}) = 15/31 = 0,483871$

Jadi, peluang Jenis Kelamin Perempuan Tidak Diare sebesar 0,483871.

Perhitungan diatas dapat dibuatkan model Naive Bayes dengan Framework RapidMiner versi 8.2. Sehingga didapatkan hasil seperti gambar 2. dibawah ini:

**Simple Distribution**

Distribution model for label attribute Hasil

Class Ya (0.544)  
7 distributions

Class Tidak (0.456)  
7 distributions

Sumber: (Data Hasil Olahan dengan Rapid Miner, 2023)

Gambar 2. Hasil Perhitungan Probabilitas dengan RapidMiner

Pada probabilitas prior terdapat dua kelas yang dibentuk, yaitu:

Kelas Prediksi = Ya Diare

Kelas Prediksi = Tidak Diare

Probabilitas prior digunakan untuk menentukan kelas pada kasus baru yang terlebih dahulu dihitung probabilitas posteriornya. Jika ada kasus baru seperti yang terlihat pada tabel dibawah ini:

Tabel 5. Perhitungan probabilitas prior kelas baru

DATA X		P(X Ci)	
Atribut	Nilai	Ya Diare	Tidak Diare
Jenis Kelamin	Perempuan	0,567568	0,483871
Usia (Bulan)	36-47	0,162162	0,096774
Frekuensi BAB	>3	0,891892	0,580645
Konsistensi Tinja	Lembek	0,324324	0,580645
Mata	Cekung	1,000000	0,000000
Keinginan Untuk Minum	Ingin minum terus	0,270270	0,580645
Keadaan Umum	Tidak sadar	0,486486	0,548387

Sumber:(Data Hasil Olahan: 2023)

Setelah diketahui probabilitas setiap atribut terhadap probabilitas tiap kelas atau  $P(X|Ci)$ , maka langkah selanjutnya adalah menghitung total keseluruhan probabilitas tiap kelas:

$$P(X|Diagnosa = Ya Diare) = P(\text{Jenis Kelamin} = \text{Perempuan} | \text{Ya Diare}) * P(\text{Usia(Bulan)} = 36-47 | \text{Ya Diare}) * P(\text{Frekuensi BAB} = >3 | \text{Ya Diare}) * P(\text{Konsistensi Tinja} = \text{Lembek} | \text{Ya Diare}) * P(\text{Mata} = \text{Cekung} | \text{Ya Diare}) * P(\text{Keinginan Untuk Minum} = \text{Ingin minum terus} | \text{Ya Diare}) * P(\text{Keadaan Umum} = \text{Tidak Sadar} | \text{Ya Diare})$$

$$P(X|Diagnosa = Ya Diare) = 0,567568 * 0,162162 * 0,891892 * 0,324324 * 1,000000 * 0,270270 * 0,486486 = 0,0035005$$

$$P(X|Diagnosa = Tidak Diare) = P(\text{Jenis Kelamin} = \text{Perempuan} | \text{Tidak Diare}) * P(\text{Usia(Bulan)} = 36-47 | \text{Tidak Diare}) * P(\text{Frekuensi BAB} = >3 | \text{Tidak Diare}) * P(\text{Konsistensi Tinja} = \text{Lembek} | \text{Tidak Diare}) * P(\text{Mata} = \text{Cekung} | \text{Tidak Diare}) * P(\text{Keinginan Untuk Minum} = \text{Ingin minum terus} | \text{Tidak Diare}) * P(\text{Keadaan Umum} = \text{Tidak Sadar} | \text{Tidak Diare})$$

$$P(X|Diagnosa = Tidak Diare) = 0,483871 * 0,096774 * 0,580645 * 0,580645 * 0,000000 * 0,580645 * 0,548387 = 0$$

Dengan menggunakan rumus Naive Bayes maka dapat dihitung:  
 $P(X|Diagnosa = Ya Diare) (\text{Ya Diare}) = 0,0035005 * 0,544118 = 0,0019047$

$$P(X|Diagnosa = Tidak Diare) (\text{Tidak Diare}) = 0 * 0,455882 = 0$$

Berdasarkan perhitungan diatas dapat diketahui bahwa kelas yang mempunyai nilai probabilitas terbesar adalah kelas Ya Diare, sehingga dapat disimpulkan contoh kasus diprediksi masuk dalam kelas Ya Diare.

1. Hasil Pengujian dengan Metode Naive Bayes

Hasil dari uji coba yang dilakukan yaitu untuk menghasilkan nilai accuracy dan nilai AUC (Area Under Curve). Hasil pengujian dengan menggunakan metode Naive Bayes untuk Prediksi Penyakit Diare Pada Balita dapat dilihat pada Gambar 3. dibawah ini:

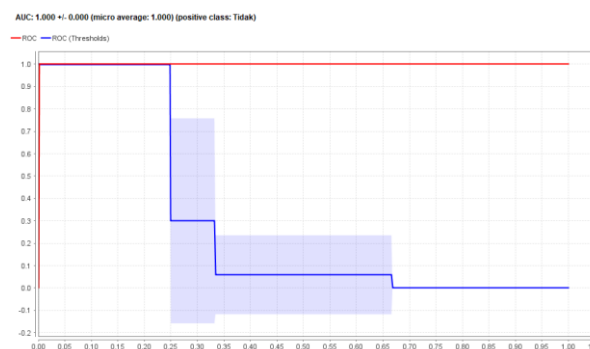
accuracy: 98.57% +/- 4.29% (micro average: 98.53%)

	true Ya	true Tidak	class precision
pred. Ya	36	0	100.00%
pred. Tidak	1	31	96.88%
class recall	97.30%	100.00%	

Sumber: (Data Hasil Olahan dengan Rapid Miner, 2023)

Gambar 3. Confusion Matrix

Dapat dilihat untuk nilai Accuracy nya sebesar 98,57%. Sedangkan hasil pengujian terhadap data testing untuk metode naive bayes terdapat nilai ROC diketahui pada Gambar 3 dibawah ini:



Sumber: (Data Hasil Olahan dengan Rapid Miner, 2023)

Gambar 4. Nilai AUC Metode Naive Bayes dalam Grafik ROC

Berdasarkan nilai AUC sebesar 1,000 yang ditunjukkan gambar diatas maka akurasi memiliki tingkat Excellent Classification.

IV. KESIMPULAN

Proses pembelajaran eksperimen ini menggunakan tools Rapid Miner. Dan proses pengujiannya menggunakan Confusion Matrix yaitu Accuracy dan ROC Curve untuk melihat tingkat keakuratan prediksi diagnosa penyakit diare pada balita. Atribut yang digunakan yaitu jenis kelamin, usia (bulan), frekuensi bab, konsistensi tinja, mata, keinginan untuk minum, keadaan umum dan satu atribut prediksi. Dari hasil penelitian nilai akurasi metode Naive Bayes digunakan untuk mendiagnosis penyakit diare pada balita dapat menghasilkan tingkat keakuratan diagnosis sebesar 98,53%, hasil tersebut di dapat dari confusion matrix. Sedangkan evaluasi menggunakan ROC Curve diperoleh hasil untuk metode naive bayes bernilai 1,000 dengan tingkat diagnosa excellent classification.

V. REFERENSI

Bahri, S., Marisa Midyanti, D., & Hidayati, R. (2018). Perbandingan Algoritma Naive Bayes dan C4.5 Untuk Klasifikasi Penyakit Anak. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 11–2018.

Fauziningrum, E., & Suryaningsih, E. I. (2021). Evaluasi Dan Prediksi Penguasaan Bahasa Inggris Maritim Menggunakan Metode Decision Tree Dan Confusion Matrix (Studi Kasus Di Universitas Maritim Amni). *Angewandte Chemie International Edition*, 6(11), 951–952., 5–24.

Giovani, A. P., Ardiansyah, A., Haryanti, T.,

- Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. *Jurnal Teknoinfo*, 14(2), 115. <https://doi.org/10.33365/jti.v14i2.679>
- Hartati, S., & Nurazila, N. (2018). Faktor Yang Mempengaruhi Kejadian Diare Pada Balita Di Wilayah Kerja Puskesmas Rejosari Pekanbaru. *Jurnal Endurance*, 3(2), 400. <https://doi.org/10.22216/jen.v3i2.2962>
- Kesehatan, K. (2023). *Laporan Kinerja 2022 Direktorat Pencegahan dan Pengendalian Penyakit Menular Kementerian Kesehatan*.
- Laila Sari, D., Saputra, M., & Gemasih, H. (2022). Penerapan Data Mining Dalam Proses Prediksi Perceraian Menggunakan Algoritma Naive Bayes Di Kabupaten Aceh Tengah. *Jurnal Teknik Informatika Dan Elektro*, 4(1), 23–35. <https://doi.org/10.55542/jurtie.v4i1.112>
- Loelianto, I., Thayf, M. S. S., & Angriani, H. (2020). Implementasi Teori Naive Bayes Dalam Klasifikasi Calon Mahasiswa Baru Stmik Kharisma Makassar. *SINTECH (Science and Information Technology) Journal*, 3(2), 110–117. <https://doi.org/10.31598/sintechjournal.v3i2.651>
- Mutiara, E.-. (2020). Algoritma Klasifikasi Naive Bayes Berbasis Particle Swarm Optimization Untuk Prediksi Penyakit Tuberculosis (Tb). *Swabumi*, 8(1), 46–58. <https://doi.org/10.31294/swabumi.v8i1.7668>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 5(2), 697–711.
- Putri, A. N. (2017). Penerapan Naive Bayesian Untuk Perankingan Kegiatan Di Fakultas Tik Universitas Semarang. *Simetris : Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 8(2), 603. <https://doi.org/10.24176/simet.v8i2.1545>
- Rahman, S., & Muzakir, A. (2020). Implementasi sistem pakar untuk menentukan penyakit diare menggunakan algoritma CBR (Case Based Reasoning). *Bina Darma Conference on ...*, 31–39. Retrieved from <https://conference.binadarma.ac.id/index.php/BDCCS/article/download/1636/670>
- Ridwan, A. (2020). Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan)*, 4(1), 15–21. <https://doi.org/10.47970/siskom-kb.v4i1.169>
- Sinum, M. B. A. (2021). Hubungan Program Open Defecation Free (ODF) oleh Pemerintah dengan Kejadian Diare. *Jurnal Medika Hutama*, 2(3), 928–933. Retrieved from <http://jurnalmedikahutama.com/index.php/JMH/article/view/191/125>
- Syarah, M. S., Wati, M., & Puspitasari, N. (2022). Klasifikasi Penderita ISPA Menggunakan Metode Naive Bayes Classifier. *Innovation in Research of Informatics (INNOVATICS)*, 4(1), 8–15. <https://doi.org/10.37058/innovatics.v4i1.4427>
- Wahyuningsih, D., & Patima, E. (2018). Penerapan Naive Bayes Untuk Penerimaan Beasiswa. *Telematika*, 11(1), 135. <https://doi.org/10.35671/telematika.v11i1.665>

Tabel 3. Sampel Data Pasien

Jenis Kelamin	Usia (bulan)	Frekuensi BAB	Konsistensi Tinja	Mata	Keinginan untuk minum	Keadaan Umum	Hasil
Perempuan	6-11	>3	Cair	Cekung	Malas minum	Gelisah	Ya
Laki-Laki	6-11	>3	Cair	Cekung	Malas minum	Gelisah	Ya
Perempuan	12-23	>3	Cair	Cekung	Malas minum	Gelisah	Ya
Laki-Laki	36-47	>3	Cair	Cekung	Malas minum	Gelisah	Ya
Perempuan	6-11	>3	Lembek	Cekung	Malas minum	Gelisah	Ya
Laki-Laki	36-47	>3	Cair	Cekung	Malas minum	Tidak sadar	Ya

Laki-Laki	24-35	>3	Cair	Cekung	Malas minum	Tidak sadar	Ya
Perempuan	48-59	>3	Cair	Cekung	Malas minum	Tidak sadar	Ya
Perempuan	24-35	>3	Cair	Cekung	Malas minum	Tidak sadar	Ya
Perempuan	6-11	>3	Cair	Cekung	Ingin minum terus	Tidak sadar	Ya
Laki-Laki	6-11	>3	Lembek	Tidak Cekung	Ingin minum terus	Tidak sadar	Tidak
Perempuan	12-23	<=3	Lembek	Tidak Cekung	Ingin minum terus	Tidak sadar	Tidak
Perempuan	12-23	>3	Padat	Tidak Cekung	Normal	Sadar	Tidak
Laki-Laki	24-35	<=3	Padat	Tidak Cekung	Normal	Sadar	Tidak
Laki-Laki	36-47	>3	Padat	Tidak Cekung	Normal	Sadar	Tidak
Laki-Laki	48-59	<=3	Lembek	Tidak Cekung	Ingin minum terus	Tidak sadar	Tidak
Laki-Laki	48-59	>3	Padat	Tidak Cekung	Normal	Sadar	Tidak
Perempuan	48-59	<=3	Lembek	Tidak Cekung	Ingin minum terus	Tidak sadar	Tidak
Perempuan	6-11	<=3	Lembek	Tidak Cekung	Ingin minum terus	Tidak sadar	Tidak
Perempuan	12-23	>3	Lembek	Tidak Cekung	Ingin minum terus	Tidak sadar	Tidak
Laki-Laki	12-23	<=3	Lembek	Tidak Cekung	Ingin minum terus	Tidak sadar	Tidak

Sumber : (Klinik Gema Medical Center Cabang Cicurug, 2023)

Hasil dari perhitungan keseluruhan atribut dapat dilihat pada tabel dibawah ini:  
Tabel 4. Probabilitas Prior

Atribut	Jumlah Kasus (S)	Ya Diare (Si)	Tidak Diare TB (Si)	P(X Ci)	
				Ya Diare	Tidak Diare
Total	68	37	31	0,544118	0,455882
Jenis Kelamin	Laki-Laki	32	16	0,432432	0,516129
	Perempuan	36	21	0,567568	0,483871
Usia (Bulan)	<6	7	5	0,135135	0,064516
	6-11	15	8	0,216216	0,225806
	12-23	12	6	0,162162	0,193548
	24-35	11	5	0,135135	0,193548
	36-47	9	6	0,162162	0,096774
Frekuensi BAB	48-59	14	7	0,189189	0,225806
	<=3	17	4	0,108108	0,419355
	>3	51	33	0,891892	0,580645
Konsistensi	Cair	25	25	0,675676	0,000000

Tinja	Lembek	30	12	18	0,324324	0,580645
	Padat	13	0	13	0,000000	0,419355
Mata	Cekung	37	37	0	1,000000	0,000000
	Tidak Cekung	31	0	31	0,000000	1,000000
Keinginan	Ingin Minum	28	10	18	0,270270	0,580645
Untuk	Terus					
Minum	Malas Minum	27	27	0	0,729730	0,000000
	Normal	13	0	13	0,000000	0,419355
Keadaan	Gelisah	18	18	0	0,486486	0,000000
Umum	Sadar	15	1	14	0,027027	0,451613
	Tidak Sadar	35	18	17	0,486486	0,548387

Sumber : (Data Hasil Olahan: 2023)