



PENERAPAN ALGORITMA NAIVE BAYES DALAM MEMPREDIKSI PENYAKIT LAMBUNG

Application of Naive Bayes Algorithm in Predicting Gastric Disease

Jefi¹, Hendri², Nurul Afni³, Agus Salim⁴, Yana Iqbal Maulana⁵
Teknologi Informasi¹, Teknik Informatika², Sistem Informasi^{3,4,5}
Fakultas Teknik & Informatika^{1,3,4,5}, Fakultas Teknologi Informasi²
Universitas Bina Sarana Informatika^{1,3,4,5}, Universitas Nusa Mandiri²,

jefi.jfi@bsi.ac.id¹, hendri.hed@nusamandiri.ac.id²,
nurul.nrf@bsi.ac.id³, agus.asm@bsi.ac.id⁴, yana.yim@bsi.ac.id⁵

Received: November 21, 2021. **Revised:** November 27, 2021. **Accepted:**
November 29, 2021. Published: December 02, 2021. **Issue Period:** Vol.5
No.2 (2021), Page 524-531

Abstrak: Lambung adalah sebuah organ pencernaan yang berfungsi sebagai tempat untuk menyimpan makanan, dan mencerna makanan menjadi ukuran yang lebih kecil yang kemudian akan diproses oleh tubuh agar menjadi energi. Penyakit pada lambung memiliki gejala yang berbeda dan ada pula yang memiliki gejala yang serupa, penyakit pada lambung adalah sebuah penyakit yang sering dialami oleh masyarakat. Penyakit ini biasanya disebabkan oleh pola makan yang kurang baik. Untuk mengetahui jenis penyakit lambung apa yang sedang di derita, pasien harus melakukan pemeriksaan diri ke rumah sakit atau ke klinik, namun karena keterbatasan waktu untuk melakukan pemeriksaan kerumah sakit atau ke klinik. Oleh karen itu diperlukan sebuah metode dan tools yang mampu memprediksi penyakit pada lambung dengan mempertimbangkan gejala yang diderita oleh pasien, sehingga akan mempermudah proses dalam memprediksi penyakit pada lambung. Algoritma Navie Bayes merupakan salah satu algoritma dari metode klasifikasi data mining yang dapat digunakan untuk memprediksi penyakit pada lambung, Naive Bayes memberikan tingkat akurasi kebenaran sebesar 75% dengan nilai area under the curve (AUC) sebesar 0,852. Dengan hal ini menunjukkan bahwa model tersebut melalui penerapan pada tools rapidminer dengan memperoleh hasil akurasi yang baik.

Kata kunci: Penyakit Lambung, Naive Bayes, Rapidminer.

Abstract: The stomach is a digestive organ that functions as a place to store food, and digest food into smaller sizes which will then be processed by the body to become energy. Diseases of the stomach have different symptoms and some have similar symptoms, gastric disease is a disease that is often experienced by the public. This disease is usually caused by a poor diet. To find out what type of gastric disease is being suffered, the patient must do a self-examination to the hospital or to the clinic,





but due to limited time to do an examination at the hospital or to the clinic. Therefore we need a method and tools that are able to predict diseases of the stomach by considering the symptoms suffered by the patient, so that it will facilitate the process of predicting diseases of the stomach. The Navie Bayes algorithm is one of the algorithms of data mining classification methods that can be used to predict gastric disease, Naive Bayes provides an accuracy rate of 75% with an area under the curve (AUC) value of 0.852. With this, it shows that the model is applied to rapidminer tools by obtaining good accuracy results.

Keywords: component; Gastric Disease, Naive Bayes, Rapidminer.

I. PENDAHULUAN

Setiap manusia tidak tertutup kemungkinan untuk terkena penyakit. Banyak faktor penyebab yang dapat memicu penyakit, misalnya pola makan, pola tidur, lingkungan, tekanan psikis dan banyak lagi (Raharjo, Damiyana, & Hidayatullah, 2016) Gastritis adalah salah satu gangguan pencernaan yang banyak diderita masyarakat dunia. Hampir 10persen penduduk dunia menderita gastritis.¹ Berdasarkan penelitian WHO, insiden gastritis di dunia telah mencapai sekitar 1,8-2,1 juta dari jumlah penduduk setiap tahunnya (Arikah & Muniroh, 2015), Penyakit pencernaan adalah penyakit yang (Ma'rifati & Kesuma, 2018)

Jenis kanker penyebab kepatian tersebut antara pria dan wanita adalah salah satunya berasal dari lambung (Wahyuningsih, 2010)

Gastritis biasanya dianggap sebagai suatu hal yang remeh namun gastritis merupakan awal dari sebuah penyakit yang dapat menyusahkan seseorang (Nurkholis, Riyantomo, & Tafrikan, 2008), Lambung merupakan organ dalam tubuh manusia yang cukup rentan terinfeksi bakteri atau terluka (Ardiansyah, Fauziah, & Ningsih, 2018).

Salah satu bakteri yang dapat menimbulkan penyakit adalah bakteri Salmonella. Salmonella adalah bakteri gram negatif, berbentuk spora yang memfermentasi glukosa menjadi Enterobacteria Penyakit tersebut dapat disebarkan melalui makanan yang terkontaminasi bakteri (Kirman, Saputra, & Sukmana, 2019)

Beberapa jenis penyakit lambung memiliki gejala yang berbeda maka dari itu ketika sedang mengalami gejala seperti penyakit lambung sebaiknya segera melakukan pemeriksaan langsung ke instansi kesehatan. Hasil penelitian sebelumnya, telah menerapkan metode naïve bayes dalam menangani permasalahan kurangnya tenaga dokter khusus menangani penyakit lambung (Wicaksono et al., 2019)

Pemeriksaan penyakit lambung dengan menerapkan sistem manual, yaitu berkonsultasi secara langsung atau tatap muka antara dokter dengan pasien memiliki kelemahan. Kelemahannya adalah penyiapan biaya yang tidak sedikit oleh pasien. Sehingga dibangun sebuah sistem pakar dengan menerapkan metode Naïve Bayes untuk mempermudah dalam melakukan konsultasi (Hermanto & Jollyta, 2021)

Stres yang berkepanjangan ini muncul karena gaya hidup saat ini yang serba cepat akibat tuntutan hidup dan tuntutan kerja, misalnya mobilitas yang tinggi maupun beban kerja yang dirasakan berat (Subekti & Utami, 2011)

Kesadaran akan kesehatan masyarakat yang masih rendah, kebiasaan hidup yang selalu ingin hidup praktis, perilaku dan pola pikir yang mengarah bergaya hidup tidak sehat, pengetahuan masyarakat yang sedikit dari gejala awal dari suatu penyakit merupakan faktor-faktor penyebab penyakit menjadi parah ketika penderita ditangani oleh tenaga paramedis (Akmal & Winiarti, 2014)

Oleh karena itu diperlukan sebuah metode dan tools yang akurat yang mampu memprediksi penyakit lambung dengan mempertimbangkan gejala yang di derita oleh pasien, sehingga metode yang akan kami terapkan ini dapat diaplikasikan dalam memprediksi penyakit lambung. Adapun tujuan dari penelitian ini adalah untuk Memudahkan dalam memprediksi jenis penyakit lambung dari gejala-gejala yang ada dengan menggunakan metode Naive Bayes, serta untuk menyediakan metode dan tools yang dapat digunakan dalam memeriksa pasien pengidap penyakit lambung.



II. METODE DAN MATERI

Adapun metode yang digunakan oleh penulis dalam penulisan penelitian ini adalah metode kualitatif. Dimana pada penelitian ini, peneliti secara langsung terjun di lapangan dalam proses penelitiannya, sehingga apabila terdapat informasi yang kurang sesuai dengan yang dibutuhkan atau di tafsirkan, maka peneliti dapat mengkonfirmasi dan melakukan pengecekan kembali pada subject. Populasi dari penelitian ini adalah pasien yang mengalami sakit pada lambung dimana penyakit tersebut terbagi menjadi 3 jenis yaitu Dispepsia, GERD dan Gastritis. Ketiga penyakit tersebut peneliti ambil dari jenis penyakit lambung yang banyak dialami oleh pasien.

2.1. Metode Pengumpulan Data

Pada penelitian ini, peneliti melakukan proses penelitian menjadi beberapa tahap. Yaitu sebagai berikut :

1. Proses Pengambilan Data
Pada proses ini peneliti melakukan proses pengambilan data, dataset yang diambil berupa data penyakit lambung diambil dari kaggle terdapat 50 data.
2. Proses Pengolahan Data.
Pada proses ini peneliti melakukan proses pengolahan data. Data yang diperoleh diolah kembali, untuk dapat mempermudah proses perhitungan.
3. Proses Perhitungan dengan Menggunakan Metode Algoritma *Naive Bayes*
Pada proses selanjutnya data yang telah di olah menghasilkan 50 data dimana data tersebut dibagi menjadi 38 data untuk data training dan 12 data untuk data testing. Selanjutnya data tersebut diolah kembali dengan melakukan proses perhitungan dengan metode algoritma *Naive Bayes*.
4. Proses Pengujian dengan *Tools RapidMiner 9.5*

Pada proses ini setelah peneliti melakukan proses perhitungan secara manual, proses selanjutnya adalah melakukan pengujian dari hasil perhitungan yang dilakukan secara manual, dengan menggunakan bantuan tools rapidminer. Yang fungsinya adalah untuk melakukan perbandingan tingkat akurasi pada diagnosa

III. PEMBAHASA DAN HASIL

3.1. Menghitung Jumlah Class atau Label

Menghitung jumlah *class* atau label adalah tahapan pertama pada perhitungan *Naive Bayes*. *Class* tersebut diantaranya yaitu; “Tepat” dan “Tidak Tepat”. Menghitung jumlah *class* dilakukan dengan cara menjumlahkan data sesuai dengan masing-masing *class*, lalu dibagi dengan keseluruhan jumlah data *training*.

$P(\text{Status_Diagnosa} = \text{“Tepat”}) = 27/38 = 71\%$.

$P(\text{Status_Diagnosa} = \text{“Tidak Tepat”}) = 11/38 = 29\%$

Berikut hasil menghitung jumlah class dalam label:

Tabel 1. Menghitung Jumlah *Class* atau *Label*

1. Menghitung Jumlah Label S_Penyakit			
Label	Jml Data	Jml Seluruh Data	Hasil
P(Benar)	27	38	0,710526316
P(Tidak)	11	38	0,289473684

Sumber: Penelitian (2021)



DOI: 10.52362/jisicom.v5i2.659

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

3.2. Menghitung Probabilitas (Peluang) Masing-masing Variabel

Tahapan dalam menghitung *probabilitas* (peluang) dilakukan dengan cara menjumlahkan *class* “Tepat” dan “Tidak Tepat” dari masing-masing variabel, lalu di bagi dengan jumlah dari masing-masing *class*.

Tabel 2. Peluang Jenis Kelamin

Peluang Jenis Kelamin				
Jenis Kelamin	Tepat	Tidak Tepat	P(Tepat)	P(T.Tepat)
Perempuan	16	4	0,59259259	0,3636364
Laki-laki	11	7	0,40740741	0,6363636
Total	27	11		

Sumber: Penelitian (2021)

Tabel 1. Peluang G1

Peluang G1				
G1	Tepat	Tidak Tepat	P(Tepat)	P(T.Tepat)
Mual	10	6	0,37037037	0,54545455
R.Mulut Asam	6	2	0,22222222	0,18181818
Nyeri Ulu Hati	9	3	0,33333333	0,27272727
Kembung	2	0	0,07407407	0
Total	27	11		

Sumber: Penelitian (2021)

Tabel 4. Peluang G2

Peluang G2				
G2	Tepat	Tidak Tepat	P(Tepat)	P(T.Tepat)
Muntah	8	6	0,2962963	0,5454545
Nyeri Dada	6	2	0,22222222	0,1818182
Mudah Kenyang	9	3	0,33333333	0,2727273
Tinja Warna Hitam	4	0	0,14814815	0
Total	27	11		

Sumber: Penelitian (2021)

Tabel 5. Peluang G3

Peluang G3				
G3	Tepat	Tidak Tepat	P(Tepat)	P(T.Tepat)
Sakit Perut	10	4	0,37037037	0,4444444
Batuk Kering	6	1	0,22222222	0,1111111
Perut Kembung	9	3	0,33333333	0,3333333
Muntah Darah	2	1	0,07407407	0,1111111
Total	27	9		

Sumber: Penelitian (2021)

Tabel 6. Peluang Diagnosa

2. Peluang Diagnosa				
Diagnosa	Tepat	Tidak Tepat	P(Tepat)	P(T.Tepat)
Gastritis	11	6	0,40740741	0,5454545
Dispepsia	9	3	0,33333333	0,2727273
GERD	7	2	0,25925926	0,1818182
Total	27	11		

Sumber: Penelitian (2021)

3.3. Mengkalikan Masing-masing Class

Proses mengkalikan masing-masing *class* “Tepat” dan “Tidak Tepat” ini bertujuan untuk mengetahui *probabilitas* yang terbesar serta digunakan untuk menentukan tepat atau tidaknya suatu diagnosa. Tahap ini dilakukan eksekusi terhadap data *testing* dengan mengkalikan masing-masing *class* “Tepat” dan “Tidak Tepat” sesuai dengan masing-masing variabel.

Tabel 7. Mengkalikan Masing-masing Class

8	PR	Nyeri Ulu Hati	Mudah Kenyang	Perut Kembung	GERD	Tepat
Tepat	0,592593	0,3333333	0,33333333	0,33333333	0,259259	0,0056902
T.Tepat	0,363636	0,2727273	0,27272727	0,33333333	0,181818	0,0016392
15	PR	Mual	Muntah	Muntah Darah	Gastritis	Tepat
Tepat	0,592593	0,3703704	0,2962963	0,07407407	0,407407	0,0019625
Tidak Tepat	0,363636	0,5454545	0,44444444	0,11111111	0,545455	0,0053427
20	PR	Mual	Muntah	Sakit Perut	Gastritis	Tidak Tepat
Tepat	0,592593	0,3703704	0,2962963	0,07407407	0,407407	0,0019625
Tidak Tepat	0,363636	0,5454545	0,44444444	0,11111111	0,545455	0,0053427
21	PR	Mual	Muntah	Sakit Perut	Gastritis	Tepat
Tepat	0,592593	0,3703704	0,2962963	0,07407407	0,407407	0,0019625
Tidak Tepat	0,363636	0,5454545	0,44444444	0,11111111	0,545455	0,0053427
22	LK	Mulut Asam	Nyeri Dada	Batuk Kering	GERD	Tepat
Tepat	0,407407	0,2222222	0,22222222	0,22222222	0,259259	0,0011591
Tidak Tepat	0,636364	0,1818182	0,18181818	0,11111111	0,181818	0,000425
27	PR	Mual	Tinja Hitam	Sakit Perut	Gastritis	Tepat
Tepat	0,592593	0,3703704	0,14814815	0,07407407	0,407407	0,0009813
Tidak Tepat	0,363636	0,5454545	0	0,11111111	0,545455	0

34	LK	Nyeri Ulu Hati	Mudah Kenyang	Perut Kembung	Dispepsia	Tepat
Tepat	0,40741	0,3333333	0,333333333	0,333333333	0,33333	0,00503
Tidak Tepat	0,63636	0,2727273	0,272727273	0,333333333	0,27273	0,004303
35	PR	Nyeri Ulu Hati	Mudah Kenyang	Perut Kembung	Dispepsia	Tepat
Tepat	0,59259	0,3333333	0,333333333	0,333333333	0,33333	0,007316
Tidak Tepat	0,36364	0,2727273	0,272727273	0,333333333	0,27273	0,002459

34	LK	Nyeri Ulu Hati	Mudah Kenyang	Perut Kembung	Dispepsia	Tepat
Tepat	0,40741	0,3333333	0,333333333	0,333333333	0,33333	0,00503
Tidak Tepat	0,63636	0,2727273	0,272727273	0,333333333	0,27273	0,004303
35	PR	Nyeri Ulu Hati	Mudah Kenyang	Perut Kembung	Dispepsia	Tepat
Tepat	0,59259	0,3333333	0,333333333	0,333333333	0,33333	0,007316
Tidak Tepat	0,36364	0,2727273	0,272727273	0,333333333	0,27273	0,002459

Sumber: Penelitian (2021)

3.4 Pengujian Tingkat Akurasi

Hasil akurasi di dapatkan dengan cara membaca hasil dari tabel .8, yaitu data yang diperoleh dari hasil pengolahan data *testing*. Dimana akan di sesuaikan jumlah masing-masing prediksi yaitu jumlah prediksi “Tepat” terhadap *class* “Tepat”, prediksi “Tidak tepat” terhadap *class* “Tidak Tepat” dan seterusnya.

Cara menghitung hasil Akurasinya adalah dengan menjumlah data yang ada pada kolom *class* “Tepat” dan prediksi “Tepat”, dan hasilnya dibagi seluruh jumlah data.

a. Akurasi Secara Manual

Di dapatkan dengan pengujian manual pada *Microsoft Excel*, dan menghasilkan akurasi dengan nilai 75 %.

Tabel 8. Hasil Akurasi

Class		
Prediksi	Tepat	Tidak Tepat
Tepat	7	2
Tidak Tepat	1	2
ACCURACY	0,75	

Sumber: Penelitian (2021)

b. Akurasi Pada Rapidminer Versi 9.5.

Hal ini bertujuan untuk mengetahui apakah hasil pengujian akurasi secara manual sudah memiliki nilai yang tepat.

Set (Apply Model) PerformanceVector (Performance)

Table View Plot View

accuracy: 75.00%

	true Tepat	true Tidak Tepat	class precision
pred. Tepat	7	1	87.50%
pred. Tidak Tepat	2	2	50.00%
class recall	77.78%	66.67%	

Sumber: Penelitian (2021)

Gambar 1. Hasil *Performance Vector* Pada Tools *Rapidminer*

Terakhir yaitu visualisasi akurasi dengan kurva ROC, dengan nilai *area under the curve* (AUC) menunjukkan nilai 0,852. Nilai tersebut berada pada kategori baik yaitu antara 0.80-0.90, yang dapat diartikan bahwa klasifikasi dengan model *naïve bayes* merupakan klasifikasi yang baik.



Sumber: Penelitian (2021)

Gambar 2. Hasil *Performance Vector* Pada Tools *Rapidminer*

IV. KESIMPULAN

Berdasarkan analisa dan perhitungan dengan menggunakan metode Naive Bayes untuk memprediksi penyakit lambung, maka dapat disimpulkan bahwa Naive Bayes memberikan tingkat akurasi kebenaran sebesar 75% dengan nilai *area under the curve* (AUC) sebesar 0,852. Dengan hal ini menunjukkan bahwa model tersebut termasuk dalam kategori klasifikasi baik karena memiliki AUC antara 0.80-0.90. serta, Algoritma Naive Bayes merupakan salah satu algoritma dari metode klasifikasi data mining yang dapat digunakan untuk memprediksi penyakit lambung, karena telah dibuktikan melalui penerapan pada tools *rapidminer* dengan memperoleh hasil akurasi yang baik.



DOI: 10.52362/jisicom.v5i2.659

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



REFERENASI

- [1] Akmal, F., & Winiarti, S. (2014). Sistem Pakar Untuk Mendiagnosa Penyakit Lambung Dengan Implementasi Metode Cbr (Case-Based. *Sarjana Teknik Informatika*, 119–129.
- [2] Ardiansyah, R., Fauziah, F., & Ningsih, A. (2018). Lambung Menggunakan Metode Dempster-Shafer, 24(3), 182–196.
- [3] Arikah, & Muniroh, L. (2015). Riwayat Makanan Yang Meningkatkan Asam Lambung Sebagai Faktor Risiko Gastritis. *Gizi Indon*, 38(1), 9–20.
- [4] Hermanto, & Jollyta, D. (2021). Penerapan Naïve Bayes Pada Sistem Pakar Diagnosa Gangguan Pencernaan Balita. *Mahasiswa Aplikasi Teknologi Komputer Dan Informasi*, 2(2), 102–106.
- [5] Kirman, Saputra, A., & Sukmana, J. (2019). Sistem pakar untuk mendiagnosis penyakit lambung dan penanganannya menggunakan metode dempster shafer, VI.
- [6] Ma'rifati, I. S., & Kesuma, C. (2018). Pengembangan Sistem Pakar Mendeteksi Penyakit Pencernaan Menggunakan Metode Naive Bayes Berbasis Web. *Evolusi*, 6(1), 41–48.
- [7] Nurkholis, A., Riyantomo, A., & Tafrikan, M. (2008). Sistem Pakar Penyakit Lambung Menggunakan Metode Forward Chaining. *Momentum*.
- [8] Raharjo, J. S. D., Damiyana, D., & Hidayatullah, M. (2016). Sistem Pakar Diagnosa Penyakit Lambung dengan Metode Forward Chaining Berbasis Android, 6(2), 1–8.
- [9] Subekti, T., & Utami, M. S. (2011). Metode Relaksasi Untuk Menurunkan Stres dan Keluhan Tukak Lambung pada Penderita Tukak Lambung Kronis, 38(2), 147–163.
- [10] Wahyuningsih, mae sri hartati. (2010). potensi pengembangan obat bahan alam indonesia untuk indonesia penyakit kanker.
- [11] Wicaksono, H., Adi, I. G., Wangiyana, S., Nizar, W. Y., Kehutanan, I., Nusa, U., & Barat, T. (2019). Studi Kolonisasi Fungi Mikoriza Arbustkuar Pada Gaharu (Gyrinops Versteegii) Dengan Sumber Inokulan Rizosfer Perkebunan Gaharu. *Agrotek*, 6(2), 45–50.

