

MODEL PREDIKSI MALARIA DENGAN ALGORITMA NAIVE BAYES

Eviyanti¹⁾, Ghiyalti Novillia²⁾, Endang Sugiarni³⁾, Maulidil Alfazi⁴⁾

^{1,2,4}Prodi Informatika Medis Fakultas Kesehatan Teknologi dan Sains Universitas Bumi Persada

³Prodi Administrasi RS, Fakultas Kesehatan Teknologi dan Sains Universitas Bumi Persada

*Correspondence: eviyanti@bumipersada.ac.id

ABSTRAK:

Latar Belakang: Migrain merupakan gangguan neurologis yang banyak dijumpai dan menimbulkan dampak signifikan terhadap kualitas hidup penderitanya. Data dari *Global Burden of Disease Study* menunjukkan bahwa migrain termasuk salah satu kondisi neurologis paling prevalen dengan tingkat disabilitas yang melebihi banyak gangguan saraf lainnya. Dalam konteks analisis data, teknik klasifikasi berperan penting untuk mengkategorikan informasi berdasarkan kriteria tertentu. Algoritma Naïve Bayes, yang mengandalkan prinsip statistik probabilistik, sering dipilih karena kemampuannya memberikan prediksi yang andal meskipun dengan data terbatas. Penelitian ini mengusulkan pemanfaatan Naïve Bayes untuk membangun sistem prediksi migrain guna meningkatkan akurasi diagnosis dan penanganan klinis. Tujuan Penelitian: Studi ini bertujuan untuk merancang model klasifikasi berbasis algoritma Naïve Bayes guna mengidentifikasi kategori migrain berdasarkan rekam medis pasien serta mengevaluasi performa prediktif model tersebut. Metode Penelitian: Penelitian ini mengaplikasikan metode *data mining* dengan algoritma Naïve Bayes sebagai klasifikator. Tahapan analisis meliputi pembangunan model, pengujian kinerja, dan evaluasi metrik akurasi. Hasil Penelitian: Model Naïve Bayes yang dibangun mencapai tingkat akurasi sebesar 80%. Temuan ini mengindikasikan bahwa pendekatan ini memiliki reliabilitas tinggi dalam mengklasifikasikan kasus migrain, sehingga berpotensi sebagai alat bantu diagnosis klinis. Kesimpulan: Hasil penelitian membuktikan bahwa Naïve Bayes merupakan algoritma efektif untuk prediksi dan klasifikasi migrain dengan presisi tinggi. Implementasi model ini dapat mendukung keputusan medis, optimasi terapi, serta menjadi landasan untuk riset neurologi berbasis kecerdasan buatan di masa depan.

Kata kunci: Migrain, Neurologi, Klasifikasi, Naive Bayes, Prediksi

ABSTRACT:

Background: Migraine is a common neurological disorder that has a significant impact on the quality of life of sufferers. Data from the Global Burden of Disease Study shows that migraine is one of the most prevalent neurological conditions with a disability rate that exceeds many other neurological disorders. In the context of data analysis, classification techniques play an important role in categorizing information based on certain criteria. The Naïve Bayes algorithm, which relies on the principle of probabilistic statistics, is often chosen because of its ability to provide reliable predictions even with limited data. This study proposes the use of Naïve Bayes to build a migraine prediction system to improve the accuracy of diagnosis and clinical treatment. Research Objectives: This study aims to design a classification model based on the Naïve Bayes algorithm to identify migraine categories based on patient medical records and evaluate the predictive performance of the model. Research Methods: This study applies a data mining method with the Naïve Bayes algorithm as a classifier. The analysis stages include model development, performance testing, and evaluation of accuracy metrics. Research Results: The Naïve Bayes model that was built achieved an accuracy level of 80%. These findings indicate that this approach has high reliability in classifying migraine cases, thus having the potential to be used as a clinical diagnostic tool. Conclusion: The results of this study prove that Naïve Bayes is an effective algorithm for predicting and classifying migraine with high precision. The implementation of this model can support medical decisions, therapy optimization, and become a foundation for future artificial intelligence-based neurological research.

Key Word: Migraine, Neurology, Classification, Naive Bayes, Prediction

PENDAHULUAN

Dalam sektor kesehatan, rekam medis memegang peran penting sebagai dasar untuk mengevaluasi perkembangan status kesehatan suatu populasi. Data ini tidak hanya berfungsi sebagai arsip, tetapi juga dapat diolah lebih lanjut untuk mengidentifikasi pola penyakit, termasuk penyakit menular seperti malaria. Analisis data rekam medis secara komprehensif memungkinkan upaya pencegahan dan pengendalian penyakit menjadi lebih terarah dan efektif.

Malaria tetap menjadi masalah kesehatan global yang signifikan, dengan 106 negara masih endemis menurut laporan World Malaria Report 2015. Upaya pengendaliannya sejalan dengan *Sustainable Development Goals* (SDGs), khususnya tujuan ketiga yang berfokus pada kehidupan sehat dan kesejahteraan masyarakat, termasuk eliminasi malaria, tuberkulosis, dan penyakit tropis terabaikan pada 2030 (Kemenkes, 2016). Komitmen ini menunjukkan pentingnya pendekatan berbasis data dalam menangani penyakit menular.

Diagnosis malaria umumnya dilakukan melalui dua metode, yaitu pemeriksaan mikroskopis dan RDT. Pemeriksaan mikroskopis melibatkan analisis sediaan darah untuk mendeteksi parasit, sedangkan RDT bekerja dengan mendeteksi antigen malaria secara imunokromatografi. Metode RDT sangat berguna di daerah terpencil atau dalam situasi wabah karena memberikan hasil cepat, meskipun memerlukan tingkat akurasi yang tinggi untuk memastikan validitas diagnosis.

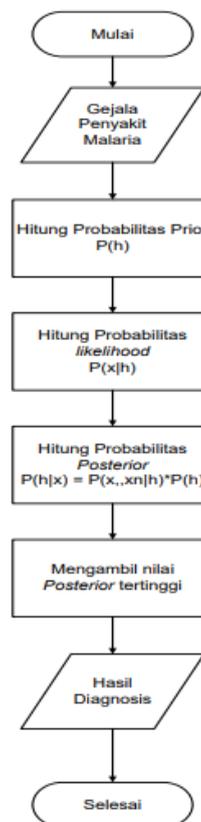
Untuk meningkatkan efektivitas pengendalian malaria, pemanfaatan teknologi informasi seperti data mining dengan algoritma Naïve Bayes dapat diterapkan pada data rekam medis. Pendekatan ini memungkinkan prediksi dini dan identifikasi kelompok berisiko, sehingga intervensi kesehatan seperti edukasi dan pencegahan dapat dilakukan secara lebih tepat sasaran.

Dengan demikian, pengolahan data kesehatan tidak hanya bersifat retrospektif, tetapi juga dapat menjadi alat strategis dalam mengurangi beban penyakit menular di masyarakat.

METODE

Penelitian ini merupakan penelitian terapan (applied research) berbasis komputasional dengan pendekatan kuantitatif-eksperimental, yang bertujuan untuk membangun dan menguji sistem deteksi penyakit malaria menggunakan algoritma Naive Bayes. Penelitian bersifat interdisipliner, menggabungkan bidang ilmu komputer dengan ilmu medis.

Penelitian ini mengembangkan pendekatan pembelajaran mesin untuk deteksi malaria menggunakan algoritma Naive Bayes. Diagram alir penelitian ini ditunjukkan pada Gambar 1.



Penelitian ini menggunakan dataset berupa rekam medis pasien

malaria yang terdiri dari 491 data dalam format Microsoft Excel, mencakup periode Agustus hingga November 2018. Data tersebut meliputi atribut seperti NIK/nomor registrasi, nama pasien, umur, jenis kelamin, alamat, tanggal kunjungan, pekerjaan, suhu badan, gejala klinis (sakit kepala, menggigil, mual/muntah, nyeri otot, berkering), serta hasil diagnosis berdasarkan jenis parasit. Untuk memfokuskan analisis, data difilter berdasarkan lokasi dan atribut yang tidak relevan seperti status kehamilan dan konfirmasi laboratorium dihilangkan. Gejala malaria yang digunakan sebagai variabel penelitian meliputi demam, sakit kepala, menggigil, mual/muntah, nyeri otot, berkering, dan hasil diagnosis.

Tahap prapemrosesan diawali dengan pembersihan data (*data cleaning*) untuk menghilangkan duplikasi, inkonsistensi, dan kesalahan pencatatan. Dari total 491 data awal, dilakukan seleksi berdasarkan wilayah studi sehingga diperoleh 118 data yang memenuhi kriteria. Sebanyak 20 data dipilih sebagai sampel representatif untuk proses validasi. Selanjutnya, dilakukan transformasi data untuk menyesuaikan format dengan kebutuhan analisis. Variabel suhu badan dikategorikan menjadi empat kelompok: Rendah ($\leq 36^{\circ}\text{-}37^{\circ}\text{C}$), Sedang ($>37^{\circ}\text{-}38^{\circ}\text{C}$), Tinggi ($39^{\circ}\text{-}40^{\circ}\text{C}$), dan Sangat Tinggi ($>40^{\circ}\text{C}$).

Data yang telah melalui tahap seleksi dan transformasi diolah menggunakan algoritma Naïve Bayes. Sebelum analisis, dataset dibagi menjadi dua subset: *data training* untuk membangun model prediksi dan *data testing* untuk menguji akurasi model. Contoh data testing meliputi variabel gejala seperti demam ($39^{\circ}\text{-}40^{\circ}\text{C}$), sakit

kepala (berat), menggigil (ya), mual/muntah (tidak), nyeri otot (sering), dan berkering (tidak), dengan kelas hasil diagnosis (positif/negatif) yang belum diketahui. Kelas hasil diagnosis dibagi menjadi dua kategori: C1 (Positif) dan C2 (Negatif).

Model prediksi diuji dengan membandingkan hasil klasifikasi dari data testing terhadap data aktual. Proses ini bertujuan untuk menilai kinerja algoritma Naïve Bayes dalam mengidentifikasi kasus malaria berdasarkan gejala klinis. Hasil evaluasi akan menunjukkan tingkat akurasi, presisi, dan recall model dalam memprediksi diagnosis.

HASIL DAN PEMBAHASAN

A. HASIL

Pengujian sistem dilakukan menggunakan 5 sampel data uji untuk mengevaluasi akurasi prediksi diagnosis malaria berbasis metode Naïve Bayes. Hasil perbandingan antara diagnosis aktual (berdasarkan data) dan diagnosis sistem disajikan pada Tabel 4.1.

Tabel 1 Hasil Pengujian Akurasi Sistem

Data Uji	Gejala (G1-G10)	Hasil Data	Hasil Sistem	Probabilitas Positif	Probabilitas Negatif	Kesesuaian
1	G1:Ya, G2:Ya, G3:Tidak, G4:Ya, G5:Ya, G6:Ya, G7:Ya, G8:Tidak, G9:Tidak, G10:Tidak	Negatif	Negatif	6.93115×10^{-1} 0	0.0001651 81	Sesuai
2	G1:Tidak, G2:Tidak, G3:Ya, G4:Ya, G5:Ya, G6:Tidak, G7:Ya, G8:Ya, G9:Tidak, G10:Tidak	Negatif	Negatif	3.25056×10^{-1} 1	0.0001416 42	Sesuai
3	G1:Ya, G2:Tidak, G3:Ya, G4:Ya, G5:Ya, G6:Tidak, G7:Ya, G8:Ya, G9:Tidak, G10:Tidak	Negatif	Negatif	4.05963×10^{-1} 0	9.34927×10^{-5}	Sesuai
4	G1:Ya, G2:Tidak, G3:Tidak, G4:Tidak, G5:Tidak, G6:Tidak, G7:Ya, G8:Tidak, G9:Tidak, G10:Ya	Negatif	Negatif	2.78178×10^{-1} 1	0.0002502 49	Sesuai
5	G1:Ya, G2:Tidak, G3:Ya, G4:Tidak, G5:Tidak, G6:Tidak, G7:Ya, G8:Tidak, G9:Ya, G10:Tidak	Positif	Negatif	5.07338×10^{-1} 0	0.0001988 01	Tidak Sesuai

Berdasarkan pengujian, sistem menghasilkan diagnosis yang sesuai dengan data aktual pada 4 dari 5 kasus (80% akurasi). Satu ketidaksesuaian

terjadi pada Data Uji 5, di mana sistem memprediksi "Negatif" padahal diagnosis sebenarnya "Positif".

B. PEMBAHASAN

Metode Naïve Bayes mencapai akurasi 80% dalam memprediksi diagnosis malaria, dengan 4 prediksi benar dan 1 kesalahan. Tingkat akurasi ini menunjukkan bahwa model cukup efektif untuk identifikasi awal, meskipun terdapat ruang untuk peningkatan. Kesalahan prediksi pada Data Uji 5 mungkin disebabkan oleh:

1. Ketidakseimbangan Data Training: Distribusi kelas (positif/negatif) yang tidak merata dapat memengaruhi probabilitas.
2. Gejala yang Tumpang Tindih: Gejala seperti demam (G7) dan nyeri otot (G5) muncul pada kedua kelas, mengurangi daya pembeda model.
3. Keterbatasan Fitur: Atribut seperti riwayat perjalanan atau kondisi lingkungan tidak dimasukkan, padahal dapat meningkatkan akurasi.

Pada semua kasus yang sesuai, probabilitas kelas "Negatif" secara konsisten lebih tinggi ($\geq 9.34927 \times 10^{-5}$) dibanding "Positif" ($\leq 6.93115 \times 10^{-10}$), mengonfirmasi kecenderungan model ke diagnosis negatif. Pada Data Uji 5, meskipun probabilitas "Negatif" (0.000198801) lebih dominan, hasil aktual "Positif" menunjukkan perlunya pemeriksaan ulang terhadap fitur yang digunakan.

Akurasi 80% mengindikasikan bahwa sistem dapat menjadi alat skrining awal di fasilitas kesehatan dengan sumber daya terbatas. Namun, untuk diagnosis definitif, disarankan:

1. Pengayaan Data: Menambahkan variabel seperti hasil

mikroskopis/RDT.

2. Optimasi Algoritma: Menerapkan teknik feature selection atau menggabungkan Naïve Bayes dengan metode lain (e.g., Random Forest).
3. Validasi dengan Sampel Lebih Besar: Pengujian dengan jumlah data lebih banyak untuk memastikan konsistensi akurasi.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa:

1. Algoritma Naïve Bayes terbukti efektif untuk memprediksi penyakit malaria dengan tingkat akurasi mencapai 80%. Hal ini menunjukkan bahwa pendekatan berbasis data mining dapat menjadi alat pendukung yang berguna dalam identifikasi awal kasus malaria.
2. Analisis gejala klinis seperti demam, sakit kepala, menggigil, dan gejala lainnya memberikan kontribusi signifikan dalam proses prediksi. Namun, terdapat ruang untuk peningkatan akurasi dengan memperluas variabel prediktor atau melakukan optimasi model.
3. Penerapan teknik data mining memungkinkan pengolahan data rekam medis pasien secara lebih sistematis, sehingga dapat menghasilkan informasi yang bermanfaat untuk pengambilan keputusan klinis dan surveilans penyakit.
4. Temuan ketidaksesuaian pada beberapa kasus (seperti pada Data Uji 5) mengindikasikan pentingnya integrasi dengan metode diagnostik lain, seperti pemeriksaan mikroskopis atau RDT, untuk hasil yang lebih andal.

DAFTAR PUSTAKA

- Darmajaya Kusri, dan Luthfi, E.M. 2009. *Algoritma Data Mining*. Yogyakarta : CV. Andi Offset
- Fatta, Al Hanif. 2007. *Analisis & Perancangan Sistem Informasi untuk Keunggulan Bersaing Perusahaan & Organisasi Modern*. Yogyakarta : CV. Andi Offset
- Hasanah, Q., Andrianto, A. dan Hidayat, M.A. 2018. *Sistem Informasi Posyandu Ibu Hamil dengan Penerapan Klasifikasi Resiko Kehamilan Menggunakan Metode Naive Bayes*
- Kusuma, P.J., 2013. *Data Mining Untuk Mendiagnosa Penyakit Infeksi Saluran Pernafasan (ISPA) Menggunakan Metode Naive Bayes*
- Mustakim, 2013. *Implementasi Data Mining Untuk Identifikasi Pola Penyakit Ispa Dengan Menggunakan Algoritma Apriori (Studi Kasus Di Uptd Puskesmas Bae Kabupaten Kudus)*
- Nurjoko, dan Kurniawan, H. 2016. *Aplikasi Datamining Untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Apriori di IIB Darmajaya Bandar Lampung*. Informatics and Bussiness Institute
- Priyanti, E., 2017. *Penerapan Algoritma Naive Bayes Untuk Klasifikasi Bakteri Gram-Negatif*, Vol. 3, No.2
- Saputra, R.A., 2014. *Komparasi Algoritma Klasifikasi Data Mining Untuk Memprediksi Penyakit Tuberculosis (Tb): Studi Kasus Puskesmas Karawang Sukabumi*
- Septiani dan Rijati, 2013. *Klasifikasi Kehamilan Risiko Tinggi Pada Ibu Hamil Puskesmas Wonotunggal Batang Menggunakan Metode Naive Bayes*

