



Penerapan Algoritma Naïve Bayes untuk Memprediksi Penyakit Malaria pada Puskesmas Hanura

Halimah*¹, Deppi Linda², Felisita Klaralia³

^{1,2,3} Jurusan Sistem Informasi, Institut Informatika Dan Bisnis Darmajaya, Jalan Zainal Abidin Pagar Alam Bandar Lampung-Lampung-Indonesia 35142

e-mail: ¹halimahyunus@darmajaya.ac.id, ²deppilinda@darmajaya.ac.id,

³klaraliaklara@gmail.com

Abstrak

Dalam bidang kesehatan data rekam medis dapat digunakan sebagai perbandingan dan tolak ukur untuk mengetahui perkembangan penyakit di suatu daerah. Namun ada baiknya data tersebut diolah menjadi data yang bermanfaat termasuk untuk dapat memprediksi suatu penyakit. Kasus malaria merupakan penyakit menular dan sangat dominan di daerah tropis/subtropis. Kabupaten Pesawaran merupakan wilayah dengan kasus malaria tertinggi dan terus meningkat setiap tahunnya. Puskesmas Hanura memiliki paling banyak penderita yang positif malaria dari beberapa wilayah kerja puskesmas yang ada di Pesawaran. Pihak Puskesmas belum memiliki sistem untuk memprediksi malaria, oleh karena itu untuk memanfaatkan teknologi sistem informasi dan untuk mencegah malaria lebih awal maka dilakukan penelitian untuk memprediksi penyakit malaria dengan metode Naïve Bayes pada Puskesmas Hanura. Dataset pasien memuat 16 atribut dan 6 diantaranya merupakan gejala penyakit malaria dengan total 118 data pasien. Hasil perhitungan dengan Naïve Bayes menunjukkan bahwa gejala penyakit yang sesuai akan menghasilkan prediksi positif. Hasil prediksi tersebut dapat digunakan untuk dugaan sementara sebelum pasien melanjutkan pemeriksaan langsung secara medis.

Kata kunci— Data mining, Naïve Bayes, Puskesmas, Prediksi, Malaria

Abstract

In the field of health, medical record data can be used as a comparison and benchmark to determine a disease progress in an area. But it is suggested that the data is processed into useful data including being able to predict a disease. Malaria is a contagious disease and is very dominant in the tropics / subtoropic regions. Pesawaran Regency is the area with the highest malaria cases that continue to increase every year. Puskesmas Hanura (Community Health Center) has the most malaria-positive sufferers from several work areas in the Public Health Center in Pesawaran District. The Puskesmas has no system to predict malaria. Therefore, to utilize information system technology and to prevent malaria early, a study was conducted to predict malaria using the Naïve Bayes method at the Hanura Community Health

Center. The patient data set contained 16 attributes and 6 of them were the symptoms of malaria with the total of 118 patient data. The results of calculations with Naïve Bayes showed that the appropriate symptoms of the disease produced positive results. The prediction results can be used for a provisional prediction before the patient continues his direct medical examination.

Keywords—3-5 Data mining, Naïve Bayes, Prediction, Puskesmas, Malaria.

1. PENDAHULUAN

Dalam bidang kesehatan, data rekam medis merupakan salah satu hal yang penting karena dapat digunakan sebagai perbandingan dan tolak ukur untuk mengetahui perkembangan kesehatan suatu daerah. Namun, ada baiknya jika data rekam tersebut dapat diolah kembali menjadi suatu data yang bermanfaat salah satunya dapat mengidentifikasi pola penyakit.

World Malaria Report 2015 menyebutkan bahwa malaria telah menyerang 106 negara di dunia. Komitmen global melalui *Sustainable Development Goals* (SDGs) berupaya melakukan pemberantasan malaria yang tertuang dalam tujuan ketiga yaitu menjamin kehidupan yang sehat dan mengupayakan kesejahteraan bagi semua orang, dengan tujuan spesifik yaitu mengakhiri epidemi AIDS, tuberkulosis, malaria, penyakit *neglected-tropical* sampai dengan tahun 2030 (Kemenkes, 2016).

Kabupaten Pesawaran merupakan kabupaten dengan kasus malaria tertinggi, hal ini dikarenakan adanya peningkatan kasus yang cukup signifikan di Pulau Pahawang yang disebabkan rusaknya lingkungan mangrove yang berakibat meluasnya tempat perindukan nyamuk menular malaria, selain itu penggunaan *Rapid Detection Test* (RDT) telah sampai pada bidan di desa sehingga kasus penemuan malaria meningkat (Dinkes Pesawaran, 2016).

Total penderita malaria di Puskesmas Hanura sekitar 68.0% pada tahun 2010, dan terus bertambah setiap tahunnya. Untuk mengidentifikasi penyakit malaria, UPT Puskesmas Hanura menggunakan 2 teknik pemeriksaan, yaitu pemeriksaan

mikroskopis, pemeriksaan ini meliputi pemeriksaan darah yang menurut teknis pembuatannya dibagi menjadi preparat darah (SDr, sediaan darah) tebal dan preparat darah tipis, untuk menentukan ada tidaknya parasit malaria dalam darah. Tes diagnostik cepat *Rapid Diagnostic Test* (RDT) adalah pemeriksaan yang dilakukan berdasarkan antigen parasit malaria dengan *imunokromatografi* dalam bentuk *dipstick*. Test ini digunakan pada waktu terjadi KLB (Kejadian Luar Biasa) atau untuk memeriksa malaria pada daerah terpencil yang tidak ada tersedia sarana laboratorium. Dibandingkan uji mikroskopis, tes ini mempunyai kelebihan yaitu hasil pengujian cepat diperoleh, akan tetapi *Rapid Diagnostic Test* (RDT) sebaiknya menggunakan tingkat *sensitivity* dan *specificity* lebih dari 95%.

Untuk memanfaatkan teknologi sistem informasi saat ini penanganan lebih lanjut terhadap pemeriksaan yang lebih efektif dengan melakukan pengolahan data rekam medis pada pasien di Puskesmas Hanura menggunakan Teknik Data Mining dengan algoritma Naïve Bayes maka data tersebut dapat digunakan untuk memprediksi penyakit malaria dan memungkinkan untuk dilakukan penyuluhan dan sosialisasi agar dapat dilakukan pencegahan lebih awal guna mengurangi peningkatan penyakit tersebut..

Beberapa *literature review* tersebut adalah sebagai berikut :

- a. Komparasi Algoritma Klasifikasi Data Mining Untuk Memprediksi Penyakit Tuberculosis (Tb): Studi Kasus Puskesmas Karawang Sukabumi yang dilakukan oleh Rizal Amegia Saputra (2014). [1]
- b. Data Mining Untuk Mendiagnosa Penyakit Infeksi Saluran Pernafasan

(ISPA) Menggunakan Metode Naive Bayes yang dilakukan oleh Pindan Jati Kusuma. [2]

- c. Klasifikasi Kehamilan Risiko Tinggi Pada Ibu Hamil Puskesmas Wonotunggal Batang Menggunakan Metode Naive Bayes yang dilakukan oleh Indri Septiani dan Nova Rijati. [3]

2. METODE PENELITIAN

Penyelesaian masalah yang digunakan oleh peneliti menggunakan metode-metode pada pengembangan sistem, berupa KDD untuk mengolah data yang akan digunakan dan SDLC sebagai metode untuk perancangan sistem yang akan dibuat.

2.1 Data Selection

Data yang digunakan adalah dataset yang berupa record yang terdiri dari 491 data penyakit malaria pada bulan Agustus sampai November 2018 dalam format Ms.Excel. Menampilkan data dengan alamat khusus Hanura saja, menghapus atribut hamil/tidak hamil, dan atribut konfirmasi lab. Memasukkan gejala umum malaria sebagai atribut terdiri dari demam, sakit kepala, menggigil, mual muntah, nyeri otot, berkering, dan *class* hasil diagnosis didapat dari jenis parasit.

2.2 Pre-processing cleaning

Sebelum melakukan proses *data mining* data perlu dilakukan pembersihan atau *data cleaning* pada *data selection*. Proses *data cleaning* mencakup antara lain duplikasi data, memeriksa data yang inkonsisten, memperbaiki kesalahan data seperti kesalahan cetak (tipografi). Pada data malaria dilakukan proses *cleaning* dengan membersihkan 491 data menjadi 118.

2.3 Transformation

Mentransformasi atau mengubah data ke dalam bentuk yang lebih sesuai untuk proses *data mining*. Data akan diubah ke dalam format yang dapat memudahkan proses untuk memprediksi penyakit malaria. Data *transformation* dapat dilihat pada tabel 2.1.

Tabel 2.1 Data Transformation

Demam	Transformasi
36°- <=37°	Rendah
37°- 38°	Sedang
39°- 40°	Tinggi
> 40°	Sangat Tinggi

2.4 Data Mining

Merupakan proses menghitung data yang sudah diseleksi, dibersihkan dan ditransform. Sebelum dihitung dengan metode *Naive Bayes*, data dibuat menjadi data training dengan atribut yang sudah disesuaikan. Data training dapat dilihat di tabel 2.2

Tabel 2.2 Data Training

NO	Demam	Sakit Kepala	Menggigil	Mual/muntah	Nyeri otot	Berkering	Class hasil
1	Rendah	Benar	Ya	Sering	Nyeri	Ya	Negatif
2	Rendah	Benar	Tidak	Tidak sering	Tidak	Ya	Negatif
3	Sedang	Benar	Tidak	Sering	Nyeri	Tidak	Negatif
4	Sedang	Ringan	Tidak	Sering	Tidak	Ya	Negatif
5	Sedang	Ringan	Ya	Tidak sering	Nyeri	Tidak	Negatif
6	Sedang	Ringan	Ya	Sering	Nyeri	Tidak	Negatif
7	Sedang	Benar	Ya	Sering	Nyeri	Tidak	Negatif
8	Sedang	Benar	Ya	Tidak sering	Tidak	Ya	Positif
9	Tinggi	Benar	Tidak	Sering	Nyeri	Tidak	Positif
10	Tinggi	Benar	Ya	Sering	Tidak	Ya	Positif
11	Tinggi	Benar	Ya	Sering	Nyeri	Ya	Positif
12	Tinggi	Ringan	Ya	Tidak sering	Nyeri	Tidak	Negatif
13	Tinggi	Ringan	Tidak	Sering	Tidak	Tidak	Negatif
14	Tinggi	Ringan	Ya	Sering	Nyeri	Tidak	Positif
15	Sangat tinggi	Benar	Ya	Tidak sering	Tidak	Ya	Positif
16	Sangat tinggi	Benar	Tidak	Tidak sering	Nyeri	Ya	Positif
17	Sangat tinggi	Benar	Tidak	Sering	Tidak	Ya	Positif
18	Sangat tinggi	Ringan	Ya	Sering	Nyeri	Tidak	Positif
19	Sangat tinggi	Ringan	Ya	Sering	Tidak	Ya	Negatif
20	Sangat tinggi	Ringan	Tidak	Sering	Nyeri	Tidak	Negatif

Total data ada 118 record, namun 20 record hanya ditampilkan sebagai sampel.

Hitung jumlah class

Jumlah dari masing-masing class hasil dibagi dengan total data yang terdapat pada data training.

$$\bullet P(\text{hasil_diagnosis}=\text{"positif"}) = \frac{9}{20} = 0.450$$

$$\bullet P(\text{hasil_diagnosis}=\text{"negatif"}) = \frac{11}{20} = 0.550$$

Hitung $P(X|C_i)$, yaitu probabilitas dari setiap atribut pada data X, kemudian dibagi dengan banyaknya jumlah class positif dan negatif :

DataX=Demam="39°-40°",Sakit Kepala="Berat", Menggigil="Ya", Mual/muntah="Tidak Sering", Nyeri Otor="Tidak Sering", Berkeringat="Ya", Class Hasil=???

- $P(\text{demam} = "39^\circ - 40^\circ" \mid \text{hasil_diagnosis} = \text{"positif"}) = 4/9 = 0,444$
- $P(\text{demam} = "39^\circ - 40^\circ" \mid \text{hasil_diagnosis} = \text{"negatif"}) = 2/11 = 0,182$
- $P(\text{sakit_kepala} = \text{"Berat"} \mid \text{hasil_diagnosis} = \text{"positif"}) = 7/9 = 0,778$
- $P(\text{sakit_kepala} = \text{"Berat"} \mid \text{hasil_diagnosis} = \text{"negatif"}) = 4/11 = 0,364$
- $P(\text{menggigil} = \text{"Ya"} \mid \text{hasil_diagnosis} = \text{"positif"}) = 6/9 = 0,667$
- $P(\text{menggigil} = \text{"Ya"} \mid \text{hasil_diagnosis} = \text{"negatif"}) = 6/11 = 0,546$
- $P(\text{mual_muntah} = \text{"Tidak Sering"} \mid \text{hasil_diagnosis} = \text{"positif"}) = 3/9 = 0,334$
- $P(\text{mual_muntah} = \text{"Tidak Sering"} \mid \text{hasil_diagnosis} = \text{"negatif"}) = 3/11 = 0,273$
- $P(\text{nyeri_otot} = \text{"Tidak"} \mid \text{hasil_diagnosis} = \text{"positif"}) = 4/9 = 0,444$
- $P(\text{nyeri_otot} = \text{"Tidak"} \mid \text{hasil_diagnosis} = \text{"negatif"}) = 4/11 = 0,364$
- $P(\text{berkeringat} = \text{"Ya"} \mid \text{hasil_diagnosis} = \text{"positif"}) = 6/9 = 0,667$
- $P(\text{berkeringat} = \text{"Ya"} \mid \text{hasil_diagnosis} = \text{"negatif"}) = 4/11 = 0,364$

Hitung total masing-masing nilai pada setiap atribut pada data X :

- $P(X \mid \text{hasil_diagnosis} = \text{"positif"}) = 0,444 \times 0,778 \times 0,667 \times 0,334 \times 0,444 \times 0,667 = 0,0228$
- $P(X \mid \text{hasil_diagnosis} = \text{"negatif"}) = 0,182 \times 0,364 \times 0,546 \times 0,273 \times 0,364 \times 0,364 = 0,0013$

Setelah didapat total dari masing-masing nilai class pada data X, kemudian kalikan dengan probabilitas class pada data training :

- $P(X \mid \text{hasil_diagnosis} = \text{"positif"}) \times P(\text{hasil_diagnosis} = \text{"positif"}) = 0,0228 \times 0,4500 = 0,0103$
- $P(X \mid \text{hasil_diagnosis} = \text{"negatif"}) \times P(\text{hasil_diagnosis} = \text{"negatif"}) = 0,0008$

2.5 Evaluation

Berdasarkan data uji yang belum diketahui class hasilnya, kemudian dihitung dengan metode Naïve Bayes maka data X = (demam= "39°- 40°", sakit kepala= "Berat", menggigil= "Ya", mual/muntah="Tidak Sering", nyeri otot= "Tidak", berkeringat= "Ya") dan class hasil yang belum diketahui, menghasilkan nilai dari hasil diagnosa positif lebih besar daripada class hasil negatif yaitu 0,0103.

3. HASIL DAN PEMBAHASAN

3.1 Tampilan Pola Naive Bayes

Menampilkan pola *Naive Bayes* dalam tabel sesuai dengan atribut. Berikut gambar 3.1

Pola Naive Bayes							
No	Demam	Sakit Kepala	Menggigil	Mual	Nyeri	Berkeringat	Hasil Class
1	36c-37c	berat	ya	tidak sering	nyeri	ya	negatif
2	36c-37c	berat	ya	tidak sering	nyeri	ya	negatif
3	36c-37c	berat	tidak	sering	nyeri	ya	negatif
4	36c-37c	berat	tidak	tidak sering	tidak	tidak	negatif
5	36c-37c	berat	tidak	tidak sering	nyeri	tidak	negatif
6	36c-37c	berat	tidak	sering	nyeri	tidak	negatif
7	36c-37c	berat	tidak	sering	tidak	ya	negatif
8	36c-37c	berat	ya	tidak sering	tidak	tidak	negatif
9	36c-37c	ringan	ya	tidak sering	nyeri	tidak	negatif
10	36c-37c	ringan	ya	tidak sering	nyeri	tidak	negatif
11	36c-37c	ringan	tidak	tidak sering	tidak	tidak	negatif
12	36c-37c	ringan	tidak	sering	tidak	ya	negatif
13	36c-37c	ringan	tidak	sering	nyeri	ya	negatif
14	36c-37c	ringan	ya	tidak sering	tidak	ya	negatif
15	36c-37c	ringan	ya	tidak sering	tidak	ya	negatif
16	36c-37c	ringan	ya	tidak sering	tidak	ya	negatif
17	36c-37c	ringan	tidak	tidak sering	nyeri	tidak	negatif
18	36c-37c	ringan	tidak	tidak sering	nyeri	tidak	negatif
19	36c-37c	ringan	tidak	tidak sering	nyeri	tidak	negatif

Gambar 3.1 Pola *Naive Bayes*

3.2 Menu Rekam Medis

Menginput gejala malaria yang diderita.
Berikut gambar 3.2

Gambar 3.2 Halaman Input Rekam Medis

3.3 Menu Hasil Prediksi Malaria

Menampilkan hasil dari prediksi dengan kesimpulanya. Berikut gambar 3.3

Data Pasien	
NIK	6
Nama Lengkap	Arel
Jenis Kelamin	Laki-Laki
Umur	15
Alamat	hanura
Pekerjaan	siswa
Hasil Class	Negatif
Kesimpulan	Negatif memiliki persentasi kemungkinan yang lebih besar yaitu: 52.885%

Gambar 3.3 Halaman hasil data pasien

3.4 Menu Cetak Data Pasien

Untuk mencetak atau *print out* table dari data pasien. Berikut gambar 3.4

NIK	Nama Lengkap	Umur	Jenis Kelamin	Alamat	Pekerjaan
1	Ahmad	12	Laki-Laki	nyaga	siswa
2	ahli	13	Laki-Laki	nyaga	siswa
3	lisa	16	Perempuan	hanura	siswa
4	agus	12	Perempuan	hanura	siswa
5	Dwi	20	Laki-Laki	hanura	siswa
6	Arel	15	Laki-Laki	hanura	siswa

Gambar 3.4 Halaman cetak data pasien SISMAL Hanura

3.5 Menu Rekam Medis

Menginput gejala malaria yang diderita.
Berikut gambar 3.5

Gambar 3.5 Halaman *input* rekam medis SISMAL Hanura

3.6 Menu Prediksi

Memprediksi gejala malaria dan dihitung dengan metode *Naive Bayes*. Klik pada data pasien yang akan di prediksi. Berikut gambar 3.6

Data Rekam Medis

[Export Excel](#)
[Cetak](#)

Caril NIK Pasien

NIK	Nama Lengkap	Umur	Jenis Kelamin	Alamat	Pekerjaan	Tanggal Kunjungi	Demam	Sakit Kepala	Menggigil	Mual	Nyeri	Berkeringat
6	Arel	15	Laki-Laki	hanura	siswa	12 September 2019	>40c	berat	tidak	sering	nyeri	tidak

Gambar 3.6 Halaman prediksi SISMAL Hanura

3.7 Menu Hasil Prediksi Malaria

Menampilkan hasil dari prediksi dengan kesimpulannya. Berikut gambar 3.7

Data Pasien

NIK	6
Nama Lengkap	Arel
Jenis Kelamin	Laki-Laki
Umur	15
Alamat	hanura
Pekerjaan	siswa
Hasil Class	Negatif
Kesimpulan	Negatif memiliki persentasi kemungkinan yang lebih besar yaitu 52.685%

Gambar 3.7 Hasil prediksi dengan kesimpulan SISMAL Hanura

Kelayakan Sistem

Kelayakan system dari implementasi data mining dengan metode *Naive Bayes* ini berisi hasil dari penerapan *Algoritma Naive Bayes* yang digunakan untuk memprediksi penyakit malaria pada Puskesmas Hanura. Pada system ini yang dapat melakukan akses adalah petugas SISMAL dan kepala Puskesmas, petugas SISMAL dapat mengakses semua menu pada system sedangkan kepala Puskesmas hanya dapat melihat pola, grafik, mencetak data-data pasien, dan mengunduh berkas dalam format excel.

4. KESIMPULAN

Berdasarkan proses Penerapan *Algoritma Naive Bayes* untuk Memprediksi Penyakit Malaria pada Puskesmas Hanura yang dimulai dari tahapan analisis sampai implementasi, maka didapat kesimpulan sebagai berikut :

1. Metode *naive bayes* dapat diterapkan untuk memprediksi penyakit malaria pada Puskesmas Hanura.
2. Penerapan metode *naive bayes* untuk memprediksi penyakit malaria dalam bentuk website dapat memberikan kemudahan bagi pihak puskesmas Hanura dalam memprediksi penyakit malaria pada pasien.
3. Dengan adanya penerapan data mining maka pihak Puskesmas dapat mengolah data pasien dan mendapatkan informasi secara mudah.

5. SARAN

Agar sistem penerapan algoritma *naive bayes* untuk memprediksi penyakit malaria pada Puskesmas Hanura dapat lebih baik, maka disarankan sebagai berikut :

1. Menambahkan informasi lebih lanjut tentang pengobatan penyakit malaria.
2. Menambahkan penyakit lain yang dapat diprediksi selain penyakit malaria pada Puskesmas Hanura.
3. Menambahkan forum untuk pasien dapat berdialog atau tanya jawab dengan dokter tentang penyakit.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Tim Redaksi Jurnal Teknik Politeknik Negeri Sriwijaya yang telah memberi kesempatan, sehingga artikel ilmiah ini dapat diterbitkan.

DAFTAR PUSTAKA

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