

PENERAPAN METODE NAÏVE BAYES PADA SISTEM PAKAR UNTUK DIAGNOSA PENYAKIT HIPERTENSI

Sarif Surejo¹, Yuni Pasmalisya Chaeriko², Pingky Septiana Ananda³

^{1,2,3} Program Studi Sistem Informasi, STMIK YMI Tegal

email: ¹sarif_surejo@yahoo.co.id, ²yunsipasmalisya712@gmail.com, ³pingkiseptiana01@gmail.com

Jl. Pendidikan No. 1 Tegal

Abstract

Keywords:

Naïve Bayes, Expert System, Hypertension

Hypertension is a condition when a person's blood pressure exceeds the normal threshold, where blood pressure is at 130/80 mmHg or more. Hypertension is not only influenced by lifestyle, but can also be influenced by gender and age. The older a person is, the higher the risk of developing hypertension. Many people don't realize they have high blood pressure, and the high blood pressure they have can be acute. People with hypertension are often lazy or don't want to go to the health service center. One alternative solution for consulting other than going to a health service center is an expert system. In this study, data analysis of hypertension was carried out using the naive Bayes method using 100 training data. The naive Bayes method will be used for the process of diagnosing hypertension by calculating the probability of each class in the training data, then calculating the prior probabilities, and calculating the final probability value to obtain the results of the diagnosis of hypertension. From the training data that has been known previously, then a test is carried out using 20 testing data that have been selected to determine the accuracy value of the naive Bayes method. Based on the test, the accuracy of the naive Bayes algorithm is 95.00%, so it can be concluded that the naive Bayes method in diagnosing hypertension provides a fast process and a method that is easy to implement and an undeniable level of accuracy.

Abstrak

Kata Kunci:

Naïve Bayes, Sistem Pakar, Hipertensi.

Hipertensi adalah suatu keadaan dimana tekanan darah seseorang melebihi ambang batas normal, dimana tekanan darah berada pada angka 130/80 mmHg atau lebih. Hipertensi tidak hanya dipengaruhi oleh gaya hidup, tetapi juga dapat dipengaruhi oleh jenis kelamin dan usia. Semakin tua seseorang, semakin tinggi risiko terkena hipertensi. Banyak orang tidak menyadari bahwa mereka memiliki tekanan darah tinggi, dan tekanan darah tinggi yang mereka miliki bisa menjadi akut. Penderita hipertensi seringkali malas atau tidak mau berobat ke pusat pelayanan kesehatan. Salah satu solusi alternatif untuk konsultasi selain pergi ke pusat pelayanan kesehatan adalah sistem pakar. Pada penelitian ini analisis data hipertensi dilakukan dengan menggunakan metode naive bayes dengan menggunakan 100 data latih. Metode naive bayes akan digunakan untuk proses diagnosa hipertensi dengan cara menghitung probabilitas setiap kelas pada data latih, kemudian menghitung probabilitas sebelumnya, dan menghitung nilai probabilitas akhir untuk mendapatkan hasil diagnosa hipertensi. Dari data latih yang telah diketahui sebelumnya, selanjutnya dilakukan pengujian dengan menggunakan 20 data pengujian yang telah dipilih untuk mengetahui nilai akurasi dari metode naive Bayes. Berdasarkan pengujian, akurasi algoritma naive Bayes adalah 95,00%, sehingga dapat disimpulkan bahwa metode naive Bayes dalam mendiagnosis hipertensi memberikan proses yang cepat dan metode yang mudah diterapkan serta tingkat akurasi yang tidak dapat disangkal.

Pendahuluan

Globalisasi merupakan peristiwa terhubungannya berbagai macam hal ke seluruh dunia. Teknologi, telekomunikasi, ilmu pengetahuan yang semakin berkembang mendukung terjadinya globalisasi. Globalisasi menyebabkan banyak perubahan dalam aspek kehidupan masyarakat. Aspek kehidupan yang mendapat pengaruh adalah kesehatan, salah satunya dalam bentuk perubahan gaya hidup masyarakat. Perubahan gaya hidup dan kebiasaan masyarakat dari tradisional ke modern yang tidak sehat rentan terkena risiko penyakit tidak menular seperti hipertensi [1].

Hipertensi adalah suatu kondisi ketika tekanan darah seseorang melebihi ambang batas normal, dimana tekanan darah berada pada 130/80 mmHg atau lebih. Hipertensi tidak hanya dipengaruhi oleh gaya hidup, tetapi juga dapat dipengaruhi oleh jenis kelamin dan usia. Semakin bertambah usia seseorang, maka risiko mengalami hipertensi semakin tinggi [2]. Di kalangan medis, hipertensi memiliki istilah silent killer yaitu penyakit yang membunuh secara diam-diam. Hal ini karena penderita hipertensi umumnya tidak menunjukkan gejala sampai tekanan darahnya sudah terlalu tinggi hingga mengancam jiwa. Namun, beberapa penderita hipertensi ada yang merasakan gejala-gejala tidak spesifik, seperti pusing, berat di tengkuk, sesak nafas, dan jantung berdebar-debar [3]. Hipertensi dapat menyebabkan beberapa komplikasi penyakit, seperti penyakit jantung, stroke, dan penyakit ginjal. Kondisi hipertensi yang berlangsung lama dan tidak segera diobati dapat berakibat fatal bagi penderitanya [2].

Sejauh ini pembahasan mengenai hipertensi sudah cukup banyak, namun penderita hipertensi justru tak kunjung turun. Banyak orang tidak menyadari bahwa mereka memiliki tekanan darah tinggi, dan tekanan darah tinggi yang mereka miliki bisa saja sudah akut. Penderita hipertensi seringkali malas atau tidak mau memeriksakan diri ke pusat pelayanan kesehatan [4]. Salah satu alternatif solusi untuk berkonsultasi selain pergi ke pusat pelayanan kesehatan adalah sistem pakar. Sistem pakar (expert system) merupakan sebuah sistem yang dapat mengadopsi pengetahuan manusia ke dalam program komputer, sehingga komputer dapat menyelesaikan masalah seperti yang dilakukan oleh seorang pakar (expert) [5]. Salah satu metode yang dapat diterapkan dalam sistem pakar adalah naïve bayes. Metode naïve bayes merupakan model probabilistik berdasarkan teorema bayes. Metode naïve bayes didefinisikan sebagai metode klasifikasi statistik. Metode ini menggunakan cara yang sederhana untuk menangani sejumlah atribut atau kelas berdasarkan teori probabilistik. Metode naïve bayes menyediakan algoritma yang mudah diimplementasikan untuk memprediksi peluang di masa depan berdasarkan pengetahuan tentang data yang telah diamati sebelumnya [6]. Pada penelitian ini, metode naïve bayes akan digunakan untuk proses diagnosa penyakit hipertensi dengan cara menghitung probabilitas dari setiap kelas pada data training, kemudian menghitung probabilitas prior, dan menghitung nilai probabilitas akhir untuk memperoleh hasil diagnosa penyakit hipertensi.

Berdasarkan permasalahan diatas maka penelitian ini dilakukan dengan tujuan untuk menerapkan metode naïve bayes dalam mendiagnosa penyakit hipertensi.

Beberapa penelitian yang telah dilakukan sebelumnya terkait penerapan metode naïve bayes pada sistem pakar untuk diagnosa penyakit hipertensi, antara lain :

Penelitian tentang Sistem Pakar Diagnosis Penyakit Hipertensi dengan Menggunakan Metode Forward Chaining dan Teori Probabilitas, penelitian ini menjelaskan diagnosa penyakit dengan melihat gejala yang tampak kemudian dari gejala tersebut dianalisis menggunakan metode forward chaining dan teora probabilitas yang hasilnya menunjukkan jenis penyakit yang diderita [7].

Penelitian tentang Implementasi Metode Pembobotan Berbasis Aturan dan Metode Profile Matching pada Sistem Pakar Medis untuk Prediksi Risiko Hipertensi, membahas tentang klasifikasi risiko hipertensi berdasarkan metode pembobotan berbasis aturan yang digunakan untuk menentukan bobot risiko berdasarkan variabel pasien dan metode profile matching digunakan untuk menghitung klasifikasi risiko berdasarkan variabel core factor dan secondary factor terhadap risiko hipertensi. Penentuan klasifikasi risiko hipertensi dilakukan dengan menghitung masing-masing kelas untuk kelas hipertensi dan kelas non-hipertensi. Hasil penelitian menunjukkan kelas yang memiliki skor tertinggi diambil sebagai hasil diagnosa [8].

Penelitian mengenai Perancangan Informasi Analisis Medik Menggunakan Logika Fuzzy Sugeno Berbasis Data Rekam Medik Pada Penyakit Hipertensi. Dalam penelitian ini membahas tentang diagnosa suatu penyakit dengan melihat gejala-gejala apa saja yang tampak pada pasien tersebut. Kemudian dari gejala-gejala tersebut di analisa menggunakan logika fuzzy (fuzzy logic) yang hasilnya menunjukkan jenis penyakit yang diderita oleh pasien [9].

Penelitian mengenai Implementasi Algoritma Naive Bayes pada Sistem Pakar untuk Diagnosis Penyakit Demam Berdarah Dengue Berbasis Website, penelitian ini menjelaskan mengenai pengujian metode naïve bayes dengan menggunakan dataset untuk menghasilkan nilai probabilitas setiap atribut dari tiap kelas yang berbeda, kemudian nilai probabilitas dari atribut tersebut di hitung untuk memprediksi penyakit [10].

Penelitian mengenai Penerapan Naive Bayes pada Sistem Pakar Diagnosa Gangguan Pencernaan Balita, penelitian ini menggunakan metode naïve bayes dengan pengujian 5 jenis gangguan dan 26 gejala. Hasil penelitian menunjukkan bahwa metode naïve bayes dapat digunakan untuk mendiagnosa gangguan saluran cerna pada balita berdasarkan gejala yang muncul terklasifikasi ke dalam Radang Usus Buntu dengan nilai prediksi yang lebih tinggi yaitu 56,71% dibandingkan dengan 2 penyakit lain [11].

Penelitian mengenai Sistem Pakar Diagnosa Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web, penelitian ini menjelaskan pengujian metode naïve bayes dengan menggunakan 5 jenis data penyakit dan 25 data gejala, kemudian dihitung dari gejala yang dipilih untuk mencari nilai probabilitas dari tiap jenis penyakit. Hasil penelitian menunjukkan jenis penyakit yang mempunyai nilai probabilitas akhir tertinggi diambil sebagai hasil diagnosa [12].

Penelitian tentang Penggunaan Metode Naïve Bayes Dalam Mengklasifikasi Pengangguran Pada Desa Bojong Kulur, penelitian ini menjelaskan metode naïve bayes dalam melakukan klasifikasi pengangguran pada Desa Bojong Kulur. Pengujian metode naïve bayes dilakukan dengan menyiapkan data training sebanyak 40 data dan data testing sebanyak 10 data yang dipilih secara acak. Hasil pengujian akurasi, presisi, dan recall klasifikasi pengangguran pada Desa Bojong Kulur dengan metode Naïve Bayes cukup tinggi yaitu sebesar 80% [13].

Penelitian selanjutnya tentang Penerapan Naive Bayes Untuk Diagnosa Penyakit Diare Usia Balita Pada Sistem Pakar Berbasis Website, penelitian ini menggunakan 140 jumlah total kasus terdiri dari 84 kasus pasien yang terjangkit diare dan 56 kasus pasien yang tidak terinfeksi penyakit diare pada anak balita diperoleh dari UPTD Puskesmas Caringin. Dari data tersebut dilakukan analisis menggunakan metode naïve bayes, kemudian diperoleh nilai probabilitas akhir tertinggi yang diambil sebagai hasil diagnosa [14].

Penelitian lain mengenai Aplikasi Sistem Pakar Untuk Mendiagnosa Penyakit Ispa Menggunakan Metode Naive Bayes Berbasis Website, penelitian ini menjelaskan aplikasi sistem pakar yang dibuat menggunakan metode naïve bayes. Terdapat data latih sebanyak 104 data dan dilakukan pengujian sebanyak 39 data uji. Dari 39 kali pengujian terdapat 36 data uji yang sesuai dan ada 3 data uji yang tidak sesuai. Hasil akurasi dari pengujian tersebut sebesar 92,3% [15].

LANDASAN TEORI

Metode Naïve Bayes

Naïve bayes merupakan merupakan model probabilistik berdasarkan *teorema bayes*. Metode *naïve bayes* didefinisikan sebagai metode klasifikasi statistik. Metode ini menggunakan cara yang sederhana untuk menangani sejumlah atribut atau kelas berdasarkan teori probabilistik. Metode *naïve bayes* termasuk ke dalam algoritma *supervised learning*. Metode *naïve bayes* menyediakan algoritma yang mudah diimplementasikan untuk memprediksi peluang di masa depan berdasarkan pengetahuan tentang data yang telah diamati sebelumnya [6].

Metode *naïve bayes* juga dinilai berpotensi baik dalam mengklasifikasikan dokumen dibandingkan dengan metode pengklasifikasian lain dalam hal akurasi dan efisiensi komputasi [10]. Persamaan dari *Teorema Bayes*:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Keterangan:

- X : Data dengan *class* yang belum diketahui
- H : Hipotesis data merupakan suatu *class* spesifik
- $P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)
- $P(H)$: Probabilitas hipotesis H (prior probabilitas)
- $P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H
- $P(X)$: Probabilitas X

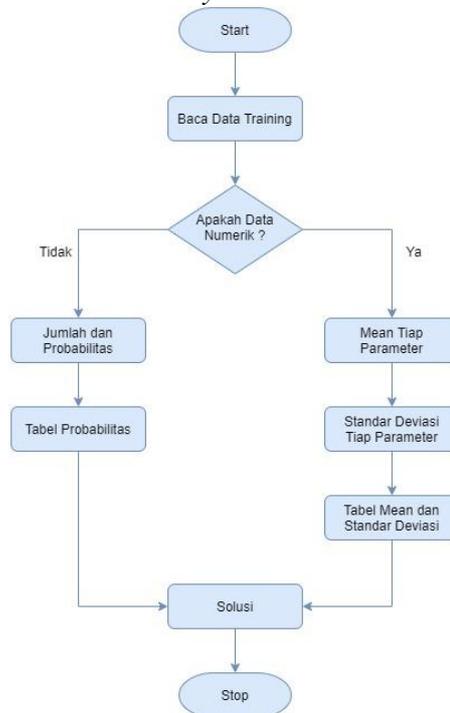
Untuk menjelaskan metode *naïve bayes*, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi *sample* yang dianalisis tersebut. Karena itu, metode *naïve bayes* di atas disesuaikan sebagai berikut:

$$P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)} \quad (2)$$

Dimana *variable C* merepresentasikan kelas, sementara *variable F1...Fn* merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya *sample* berkarakteristik tertentu kedalam kelas *C (Posterior)* adalah peluang munculnya kelas *C* (sebelum masuk *sample* tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik-karakteristik *sample* pada kelas *C* (disebut juga *likelihood*), dibagi dengan kemunculan karakteristik-karakteristik *sample* secara global (disebut juga *evidence*). Karena itu rumus diatas dapat ditulis secara sederhana sebagai berikut:

$$Posterior = \frac{prior \times likelihood}{evidence} \quad (3)$$

Berikut alur algoritma dari metode *naïve bayes*:



Gambar 2. Alur Metode *Naïve Baiyes*

Adapun keterangan dari gambar di atas sebagai berikut:

- 1). Baca data *training*
- 2). Hitung jumlah dan probabilitas, namun apabila data numerik maka:
 - a). Cari nilai *mean* dan standar deviasi dari masing-masing parameter yang merupakan data numerik. Adapun persamaan yang digunakan untuk menghitung nilai rata-rata hitung (*mean*) dapat dilihat sebagai berikut:

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (4)$$

atau

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad (5)$$

Di mana:

μ : rata-rata hitung (*mean*)

x_i : nilai *sample* ke i
 n : jumlah *sample*

dan persamaan untuk menghitung simpangan baku (standar deviasi) dapat dilihat sebagai berikut:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}} \quad (6)$$

Dimana:

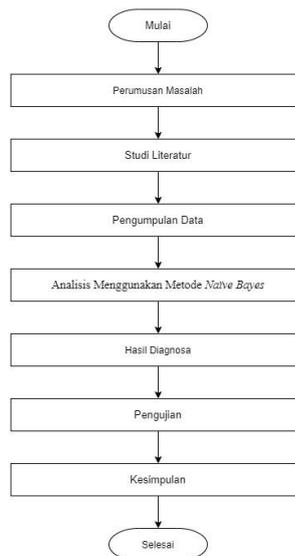
σ : standar deviasi
 x_i : nilai x ke $-i$
 μ : rata-rata hitung
 n : jumlah *sample*

- b). Cari nilai probabilistik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
- 3). Mendapatkan nilai dalam tabel *mean*, standard deviasi dan probabilitas.
- 4). Solusi kemudian dihasilkan.

Metode Penelitian

1. Tahapan Penelitian

Tahapan penelitian merupakan rangkaian bagan yang menggambarkan proses penelitian dalam mencapai tujuan penelitian. Berikut adalah tahapan penelitian yang digunakan dalam penelitian ini.



Gambar 1. Tahapan Penelitian

Berdasarkan tahapan penelitian tersebut maka masing-masing langkah dapat diuraikan sebagai berikut:

1. **Perumusan Masalah**
 Merumuskan masalah merupakan suatu cara untuk menyelesaikan atau menjawab suatu masalah, untuk merumuskan masalah pada penelitian ini maka perlu ditentukan ruang lingkup masalah yang akan diteliti.
2. **Studi Literatur**
 Pada tahap ini peneliti melakukan kajian literatur untuk mendapatkan informasi melalui buku, jurnal, dan penelitian yang telah dilakukan sebelumnya untuk mendukung penelitian.
3. **Pengumpulan Data**
 Pada tahap pengumpulan data peneliti menggunakan data *primer* dan data *sekunder*. Data *primer* didapat dengan cara tanya jawab langsung kepada Dokter di Klinik Pratama Mitra Sehat untuk

mendapatkan informasi yang berkaitan dengan penyakit hipertensi, sedangkan untuk data *sekunder* berupa data rekam medis pasien Klinik Pratama Mitra Sehat.

4. Analisis Menggunakan Metode *Naïve Bayes*
 Pada tahap ini melakukan analisis data untuk menyesuaikan proses data yang akan diolah pada metode *naïve bayes* dengan menggunakan data *training*.
5. Hasil Diagnosa
 Pada tahap ini mendapatkan hasil identifikasi dari penerapan algoritma *naïve bayes*.
6. Pengujian
 Tahap pengujian dilakukan untuk mengetahui nilai akurasi metode *naïve bayes* dalam mendiagnosa penyakit hipertensi.
7. Kesimpulan
 Bagian akhir yang menjelaskan tentang keseluruhan atau inti dari suatu gagasan atau penelitian.

Hasil dan Pembahasan

1. Analisis Menggunakan Metode *Naïve Bayes*

Data *training* merupakan data pembelajaran untuk memprediksi peluang dari data yang sebelumnya. Data *training* dalam penelitian ini diambil dari data rekam medis Klinik Pratama Mitra Sehat. Data yang didapat peneliti sebanyak 100 data terdiri dari 58 data pasien yang menderita penyakit hipertensi dan 42 data pasien yang tidak menderita hipertensi. Untuk keperluan penelitian, dari 100 data yang terkumpul hanya ditampilkan 20 data seperti pada tabel 1.

Tabel 1. Data Pasien Klinik Mitra Sehat

No.	Jenis Kelamin	Usia	Pusing	Berat di Tengku	Sesak Nafas	Jantung Berdebar	Tekanan Darah Sistolik	Tekanan Darah Diastolik	Hasil
1	Laki-Laki	Remaja	Tidak	Tidak	Tidak	Tidak	100 mmHg	60 mmHg	Tidak
2	Perempuan	Remaja	Ya	Tidak	Tidak	Tidak	100 mmHg	60 mmHg	Tidak
3	Laki-Laki	Dewasa	Tidak	Tidak	Ya	Tidak	110 mmHg	70 mmHg	Tidak
4	Perempuan	Dewasa	Tidak	Tidak	Tidak	Tidak	110 mmHg	70 mmHg	Tidak
5	Laki-Laki	Lansia	Ya	Tidak	Tidak	Tidak	120 mmHg	70 mmHg	Tidak
6	Perempuan	Lansia	Tidak	Ya	Tidak	Tidak	120 mmHg	70 mmHg	Tidak
7	Laki-Laki	Dewasa	Ya	Ya	Tidak	Ya	130 mmHg	80 mmHg	Ya
8	Perempuan	Dewasa	Ya	Ya	Ya	Tidak	130 mmHg	80 mmHg	Ya
9	Laki-Laki	Dewasa	Ya	Ya	Ya	Ya	140 mmHg	90 mmHg	Ya
10	Perempuan	Dewasa	Ya	Ya	Tidak	Tidak	140 mmHg	90 mmHg	Ya
11	Laki-Laki	Lansia	Ya	Ya	Ya	Ya	150 mmHg	100 mmHg	Ya
12	Perempuan	Lansia	Ya	Ya	Ya	Tidak	150 mmHg	90 mmHg	Ya
13	Laki-Laki	Lansia	Ya	Ya	Ya	Tidak	160 mmHg	90 mmHg	Ya
14	Perempuan	Lansia	Ya	Ya	Tidak	Ya	160 mmHg	90 mmHg	Ya
15	Laki-Laki	Dewasa	Ya	Tidak	Ya	Ya	170 mmHg	90 mmHg	Ya
16	Perempuan	Lansia	Ya	Ya	Ya	Ya	170 mmHg	90 mmHg	Ya
17	Laki-Laki	Lansia	Ya	Ya	Tidak	Ya	180 mmHg	90 mmHg	Ya
...
98	Perempuan	Lansia	Ya	Ya	Ya	Ya	180 mmHg	90 mmHg	Ya
99	Perempuan	Lansia	Ya	Ya	Ya	Ya	190 mmHg	100 mmHg	Ya
100	Laki-Laki	Lansia	Ya	Ya	Ya	Ya	200 mmHg	100 mmHg	Ya

Dalam membuat model *naïve bayes* terlebih dahulu kita mencari probabilitas hipotesis untuk masing-masing Kelas P (H). Hipotesis yang ada yaitu pasien yang menderita penyakit hipertensi dan pasien yang tidak menderita penyakit hipertensi. Berikut perhitungan probabilitas seperti tabel 2 di bawah ini:

Tabel 2. Probabilitas Kelas

		Kelas	
		Ya	Tidak
Ya	58	Tidak	42

		Kelas	
		Ya	Tidak
P (Ya)			P (Tidak)
$= \frac{58}{100}$	0,58	$= \frac{42}{100}$	0,42

Setelah probabilitas untuk tiap hipotesis diketahui, langkah selanjutnya adalah menghitung probabilitas kondisi tertentu (probabilitas X) berdasarkan probabilitas tiap hipotesis (probabilitas H) atau dinamakan probabilitas *prior*. Hasil perhitungan probabilitas *prior* dengan menggunakan *naïve bayes* dapat dilihat pada tabel 3 di bawah ini:

Tabel 3. Probabilitas *Prior*

Atribut	Jumlah Kasus	Ya	Tidak	P(X H)		
				Ya	Tidak	
Total	100	58	42			
Jenis	Laki-Laki	48	35	13	0,60	0,31
Kelamin	Perempuan	52	23	29	0,40	0,69
	Remaja	18	0	18	0	0,43
Usia	Dewasa	30	16	14	0,28	0,33
	Lansia	52	42	10	0,72	0,24
Pusing	Ya	70	52	18	0,90	0,43
	Tidak	30	6	24	0,10	0,57
Berat di Tengkok	Ya	53	48	5	0,83	0,12
	Tidak	47	10	37	0,17	0,88
Sesak Nafas	Ya	61	47	14	0,81	0,33
	Tidak	39	11	28	0,19	0,67
Jantung Berdebar	Ya	51	46	5	0,79	0,12
	Tidak	49	12	37	0,21	0,88
Tekanan Darah Sistolik	100 mmHg	8	0	8	0	0,19
	110 mmHg	13	0	13	0	0,31
	115 mmHg	2	0	2	0	0,05
	120 mmHg	18	0	18	0	0,43
	125 mmHg	1	0	1	0	0,02
	130 mmHg	7	7	0	0,12	0
	140 mmHg	15	15	0	0,26	0
	150 mmHg	8	8	0	0,14	0
	160 mmHg	7	7	0	0,12	0
	170 mmHg	9	9	0	0,16	0
Tekanan Darah Diastolik	180 mmHg	9	9	0	0,16	0
	190 mmHg	2	2	0	0,03	0
	200 mmHg	1	1	0	0,02	0
	60 mmHg	6	0	6	0	0,14
Tekanan Darah Diastolik	70 mmHg	26	0	26	0	0,62
	75 mmHg	1	0	1	0	0,02
	80 mmHg	32	23	9	0,40	0,21
	90 mmHg	24	24	0	0,41	0
	100 mmHg	11	11	0	0,19	0

Pada tabel probabilitas *prior* terdapat dua kelas yang dibentuk, yaitu:

Kelas Hasil = Ya

Kelas Hasil = Tidak

Tabel probabilitas *prior* akan digunakan untuk menentukan kelas pada data *testing*.

Terdapat contoh data *testing* dengan data sebagai berikut:

Data laki-laki berusia lansia dengan gejala pusing (ya), berat di tengkok (ya), sesak nafas (ya), jantung berdebar (ya), tekanan darah sistolik 170 mmHg dan tekanan darah diastolik 90 mmHg. Apakah orang tersebut menderita penyakit hipertensi?

Setelah diketahui probabilitas setiap atribut terhadap probabilitas tiap kelas atau P(X|H), maka langkah selanjutnya adalah menghitung nilai probabilitas akhir untuk setiap kelas

$$\begin{aligned}
 P(X|Ya) &= P(\text{Jenis kelamin} = \text{Laki-Laki}|Ya) * P(\text{Usia} = \text{Lansia}|Ya) * P(\text{Pusing} = Ya|Ya) * P(\text{Berat di Tengkrak} = Ya|Ya) * P(\text{Sesak Nafas} = Ya|Ya) * P(\text{Jantung Berdebar} = Ya|Ya) * P(\text{Tekanan darah sistolik} = 170 \text{ mmHg}|Ya) * P(\text{Tekanan darah diastolik} = 90 \text{ mmHg}|Ya) \\
 &= \frac{35}{58} * \frac{42}{58} * \frac{52}{58} * \frac{48}{58} * \frac{47}{58} * \frac{46}{58} * \frac{9}{58} * \frac{24}{58} \\
 &= 0,60 * 0,72 * 0,90 * 0,83 * 0,81 * 0,79 * 0,16 \\
 &\quad * 0,41 \\
 &= 0,013546288
 \end{aligned}$$

$$\begin{aligned}
 P(X|\text{Tidak}) &= P(\text{Jenis kelamin} = \text{Laki-Laki}|\text{Tidak}) * P(\text{Usia} = \text{Lansia}|\text{Tidak}) * P(\text{Pusing} = Ya|\text{Tidak}) * P(\text{Berat di Tengkrak} = Ya|\text{Tidak}) * P(\text{Sesak Nafas} = Ya|\text{Tidak}) * P(\text{Jantung Berdebar} = Ya|\text{Tidak}) * P(\text{Tekanan darah sistolik} = 170 \text{ mmHg}|\text{Tidak}) * P(\text{Tekanan darah diastolik} = 90 \text{ mmHg}|\text{Tidak}) \\
 &= \frac{13}{42} * \frac{10}{42} * \frac{18}{42} * \frac{5}{42} * \frac{14}{42} * \frac{5}{42} * \frac{0}{42} * \frac{0}{42} \\
 &= 0,31 * 0,24 * 0,43 * 0,12 * 0,33 * 0,12 * 0 * 0 \\
 &= 0
 \end{aligned}$$

2. Hasil Diagnosa

Selanjutnya nilai tersebut dimasukkan untuk mendapatkan probabilitas akhir.

$$\begin{aligned}
 P(X|\text{Hasil}=Ya) * P(Ya) \\
 &= 0,58 * 0,013546288 \\
 &= 0,007856847
 \end{aligned}$$

$$\begin{aligned}
 P(X|\text{Hasil}=\text{Tidak}) * P(\text{Tidak}) \\
 &= 0,42 * 0 \\
 &= 0
 \end{aligned}$$

Hasil dari nilai probabilitas akhir terbesar berada di kelas Ya, maka orang tersebut mengalami penyakit hipertensi. Dari perhitungan inilah kita dapat mengetahui apakah pasien tersebut mengalami penyakit hipertensi atau tidak.

3. Pengujian

Pengujian dilakukan untuk memutuskan apabila perhitungan yang telah dilakukan diatas tepat untuk memutuskan status hipertensi dengan menggunakan metode *naïve bayes*. Uji coba dilakukan dengan menggunakan 20 data *testing* yang telah dipilih, kemudian data *testing* dibandingkan dengan data *training* yang telah diketahui sebelumnya. Hasil dari uji coba yang dilakukan menghasilkan nilai akurasi dan nilai *AUC (Area Under Curve)* menggunakan *Confusion Matrix* dan kurva *ROC* dari *RapidMiner*.

1. Confusion Matrix

Confusion matrix akan membentuk *matrix* yang terdiri dari *true positive* atau tupel positif dan *true negative* atau tupel negatif, kemudian masukan data *testing* yang sudah disiapkan ke dalam *confusion matrix* sehingga didapatkan nilai akurasi pada gambar 3 di bawah ini:

accuracy: 95.00%

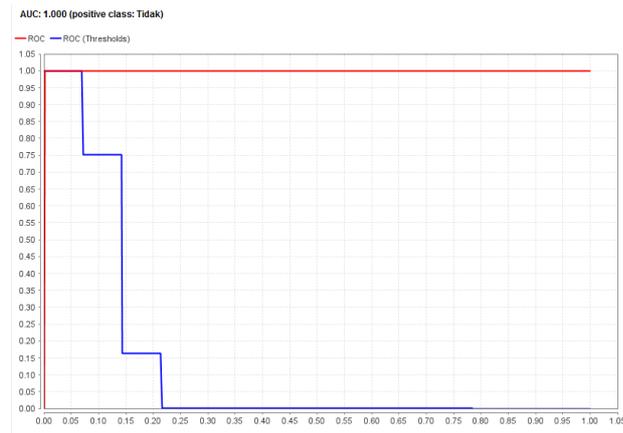
	true Ya	true Tidak	class precision
pred. Ya	13	0	100.00%
pred. Tidak	1	6	85.71%
class recall	92.86%	100.00%	

Gambar 3. Hasil Akurasi Data *Testing*

Berdasarkan hasil pengujian, nilai akurasi yang diperoleh sebesar 95.00%. Terdapat rincian jumlah *True Positive* (TP) yaitu 13, *False Positive* (FP) yaitu 0, *False Negative* (FN) yaitu 1 dan *True Negative* (TN) yaitu 6.

2. Kurva ROC

Hasil pengujian terhadap data yang digunakan untuk algoritma *naïve bayes* terhadap kurva *ROC* diketahui pada gambar 4 di bawah ini:



Gambar 4. Hasil Kurva ROC

Berdasarkan nilai *AUC* yang didapatkan dari pengujian kurva *ROC* sebesar 1.000, sehingga dari hasil tersebut menunjukkan klasifikasi yang dihasilkan termasuk kedalam tingkat *Excellent Classification*.

Kesimpulan dan Saran

Berdasarkan hasil penelitian tentang mendiagnosa penyakit hipertensi dapat ditarik beberapa kesimpulan sebagai berikut:

Metode *naïve bayes* memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk *class* yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi penyakit hipertensi berdasarkan proses klasifikasi yang dilakukan oleh metode *naïve bayes* itu sendiri.

Berdasarkan pengujian diperoleh nilai akurasi algoritma *naïve bayes* sebesar 95.00%, sedangkan evaluasi menggunakan kurva *ROC* diperoleh hasil untuk algoritma *naïve bayes* bernilai 1.000 dengan tingkat diagnosa *Excellent Classification*.

REFERENSI

- [1] U. Soleha, "Analisis Life Style Berdasarkan Faktor Karakteristik Penderita Hipertensi Di Rw. 09 Desa Sruni Gedangan Sidoarjo Jawa Timur Indonesia," *J. Heal. Sci.*, vol. 12, no. 1, pp. 93–101, 2019.
- [2] N. Widjaya, F. Anwar, R. L. Sabrina, R. R. Puspawati, and E. Wijayanti, "Hubungan Usia Dengan Kejadian Hipertensi di Kecamatan Kresik dan Tegal Angus, Kabupaten Tangerang," *J. Kedokt. Yars.*, vol. 26, no. 3, pp. 131–138, 2018, doi: 10.33476/jky.v26i3.756.
- [3] S. Susilawati, S. Nurcahyati, and A. Syaripudin, "Pengaruh Terapi Murotal Al-Quran Terhadap Stres Pasien Hipertensi Di Ruang Penyakit Dalam Rsud Waled Cirebon," *J. Kesehat. Mahardika*, vol. 6, no. 1, pp. 32–36, 2019.
- [4] D. Yolanda, Derisma, and D. Yendri, "Penerapan Metode Certainty Factor Dalam Sistem Pendeteksi Risiko Hipertensi Berbasis Smartphone," *InComTech J. Telekomun. dan Komput.*, vol. 11, no. 1, pp. 37–50, 2021.
- [5] Yuliyana and A. S. R. M. Sinaga, "Sistem Pakar Diagnosa Penyakit Gigi Menggunakan Metode Naive Bayes," *J. Fountain Informatics*, vol. 4, no. 1, pp. 19–23, 2019.
- [6] B. O. Afeni, T. I. Aruleba, and I. A. Oloyede, "Hypertension Prediction System Using Naive Bayes Classifier," *J. Adv. Math. Comput. Sci.*, vol. 24, no. 2, pp. 1–11, 2017.
- [7] E. R. Simarmata, "Sistem Pakar Diagnosis Penyakit Hipertensi Dengan Menggunakan Metode Forward Chaining Dan Toeri Probabilitas," *J. Method.*, vol. 7, no. 1, pp. 56–61, 2021.
- [8] A. Wantoro, A. Syarif, K. N. Berawi, K. Muludi, S. R. Sulistiyanti, and Sutyarso, "Implementasi Metode Pembobotan Berbasis Aturan Dan Metode Profile Matching Pada Sistem Pakar Medis Untuk Prediksi Risiko Hipertensi," *J. TEKNOINFO*, vol. 15, no. 2, pp. 134–145, 2021.
- [9] N. T. Pradipta, Fauziah, and U. Darusalam, "Perancangan Sistem Informasi Analisis Medik Menggunakan

- Logika Fuzzy Sugeno Berbasis Data Rekam Medik Pada Penyakit Hipertensi,” *J. Ilm. Inform.*, vol. 2, no. 1, pp. 59–67, 2017.
- [10] E. Nurlelah, A. Abdilah, and M. A. Ghani, “Implementasi Algoritma Naive Bayes Pada Sistem Pakar Untuk Diagnosis Penyakit Demam Berdarah Dengue Berbasis Website,” *J. Speed – Sentra Penelit. Eng. dan Edukasi*, vol. 11, no. 3, pp. 1–8, 2019.
- [11] Hermanto and D. Jollyta, “Penerapan Naive Bayes Pada Sistem Pakar Diagnosa Gangguan Pencernaan Balita,” *J. Mhs. Apl. Teknol. Komput. dan Inf.*, vol. 2, no. 2, pp. 102–106, 2020.
- [12] M. R. Handoko and Neneng, “Sistem Pakar Diagnosa Penyakit Selama Kehamilan Menggunakan Metode Naive Bayes Berbasis Web,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 50–58, 2021.
- [13] Y. S. T. Allo, V. Sofica, N. Hasan, and M. Septiani, “Penggunaan Metode Naive Bayes Dalam Mengklasifikasi Pengangguran Pada Desa Bojong Kulur,” *Bianglala Inform.*, vol. 10, no. 1, pp. 30–35, 2022.
- [14] E. Nurlelah and R. Wajhillah, “Penerapan Naive Bayes Untuk Diagnosa Penyakit Diare Usia Balita Pada Sistem Pakar Berbasis Website,” *Semin. Nas. Ilmu Pengetah. dan Teknol. Komput. Nusa Mandiri*, pp. 52–59, 2016.
- [15] F. Ramadhana, Fauziah, and Winarsih, “Aplikasi Sistem Pakar Untuk Mendiagnosa Penyakit Ispa Menggunakan Metode Naive Bayes Berbasis Website,” *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 4, no. 3, pp. 320–329, 2020.