

---

## Perbandingan Algoritma *Naive Bayes* dan C4.5 dalam Memprediksi Penyakit

**Bustamin\*<sup>1</sup>, Andriani\*<sup>2</sup>, Sugiarto Cokrowibowo<sup>3</sup>**

<sup>1</sup>STMIK Hasan Sulur Wonomulyo <sup>2,3</sup>Universitas Sulawesi Barat, Majene

Email: \*<sup>1</sup>bustaminstmik@gmail.com, <sup>3</sup>[andriani@gmail.com](mailto:andriani@gmail.com) <sup>3</sup>[sugiarto.cokrowibowo@unsulbar.ac.id](mailto:sugiarto.cokrowibowo@unsulbar.ac.id),

### **Abstrak**

Jumlah penderita diabetes setiap tahunnya selalu mengalami peningkatan. Data yang ditunjukkan oleh World Health Organization (WHO) menunjukkan terdapat sekitar 347 juta orang dari seluruh dunia yang menderita diabetes melitus, selain itu diperkirakan pula kematian yang dikarenakan diabetes terus meningkat sebesar dua pertiga kali pada tahun 2008 hingga 2030. Peningkatan jumlah penderita diabetes selain disebabkan oleh keterlambatan dalam memprediksi terjadinya, juga dikarenakan tidak sehatnya pola hidup yang dijalani. Algoritma Naive bayes dan C4.5 memiliki konsep yang didasarkan pada bahasa alami dan sangat fleksibel terhadap data-data yang kurang tepat. Semua hal ini mendasari adanya kebutuhan terhadap suatu sistem yang berfungsi sebagai alat bantu dalam menentukan bahwa pasien adalah penderita diabetes melitus dengan memanfaatkan konsep perbandingan algoritma Naive Bayes dan C4.5. Berdasarkan permasalahan diatas, dapat dikembangkan sebuah teknik datamining untuk memprediksi pasien yang teridentifikasi penyakit diabetes berdasarkan data pasien dengan mengaplikasikan Algoritma C4.5 dan algoritma Naive Bayes. Data yang telah diolah menggunakan teknik data mining tersebut mampu menghasilkan informasi guna memprediksi pasien yang teridentifikasi menderita penyakit diabetes.

**Kata kunci** — Prediksi penyakit, Penyakit diabetes, Naive Bayes, C45

### **Abstract**

The number of diabetics every year is always increasing. Data shown by the World Health Organization (WHO) shows that there are around 347 million people from all over the world who suffer from diabetes mellitus, besides that it is also estimated that deaths caused by diabetes will continue to increase by two-thirds of the years from 2008 to 2030. by delays in predicting its occurrence, also due to the unhealthy lifestyle that is lived. Naive Bayes and C4.5 have concepts that are based on natural language and are very flexible towards imprecise data. All of this is based on the need for a system that functions as an aid in determining that a patient has diabetes mellitus by utilizing the concept of comparison of the Naive Bayes algorithm and C4.5. Based on the above problems, a datamining technique can be developed to predict patients who are identified with diabetes based on patient data by applying the C4.5 Algorithm and the Naive Bayes algorithm. The data that has been processed using data mining techniques is capable of producing information to predict patient identified as having diabetes.

**Keywords** — Prediction of disease, Diabetes, Naive Bayes, C45

## 1. PENDAHULUAN

*National Institute of Diabetes and Digestive and Kidney Diseases Health Information Center* menjelaskan bahwa diabetes adalah penyakit yang dapat terjadi jika kandungan gula yang terlalu tinggi dalam darah. Tingginya kadar gula dalam darah dapat menyebabkan masalah kesehatan seiring dengan berjalannya waktu[1].

Penderita diabetes selalu meningkat setiap tahunnya. Data dari *World Health Organization* (WHO) menunjukkan bahwa sekitar 347 juta orang dari seluruh dunia menderita diabetes melitus. Kematian yang dikarenakan oleh diabetes diperkirakan terus meningkat sebesar dua pertiga kali pada tahun 2008 hingga 2030. Secara global penderita diabetes melitus terus meningkat, namun hal ini secara khusus terjadi pada negara berkembang[2].

Jumlah penderita diabetes yang meningkat disebabkan oleh pola hidup yang tidak sehat dan keterlambatan dalam meprediksi pemprediksi hal ini. Algoritma *Naive bayes* dan *C4.5* memiliki konsep yang didasarkan pada bahasa alami dan sangat fleksibel terhadap data-data yang kurang tepat [3]. Oleh sebab itu, diperlukan sebuah sistem dapat difungsikan sebagai alat untuk membantu dalam menentukan pasien penderita diabetes melitus dengan memanfaatkan konsep perbandingan algoritma *Naive Bayes* dan *C4.5*.

Algoritma *C4.5* merupakan salah satu algoritma untuk pengklasifikasian data dengan teknik *decision tree* yang terkenal dan banyak disukai karena mempunyai beberapa kelebihan. Kelebihan ini seperti dapat mengolah data numerik (kontinyu) maupun diskrit, mampu mengatasi nilai atribut yang hilang, serta menghasilkan aturan yang dapat diinterpretasikan dengan mudah dan merupakan algoritma tercepat diantara algoritma lainnya[4].

*Naive Bayes Classification* merupakan metode klasifikasi sederhana yang memanfaatkan konsep peluang, dengan asumsi bahwa setiap atribut contoh memiliki sifat saling lepas antara satu sama lain sesuai dengan atribut kelas. Algoritma klasifikasi *Naive Bayes* banyak dimanfaatkan dalam pengklasifikasian dokumen, mendeteksi spam, atau berbagai masalah lainnya terkait klasifikasi[3]. Pada penelitian ini, masalah klasifikasi yang dimaksud terkait dengan prediksi penyakit diabetes berdasarkan data gejala yang ditunjukkan oleh pasien diabetes tersebut.

Latarbelakang di atas menjadi dasar pemanfaatan suatu teknik data mining untuk memberikan prediksi pasien yang teridentifikasi penyakit diabetes dengan memanfaatkan algoritma *C4.5* dan algoritma *Naive Bayes*. Data yang telah diolah menggunakan teknik data mining tersebut diharapkan mampu menghasilkan informasi guna memprediksi pasien yang teridentifikasi menderita penyakit diabetes.

## 2. METODE

### 2.1. Penyakit diabetes

Diabetes dikategorikan sebagai penyakit kronis yang serius dikarenakan ketidakmampuan pankreas dalam memproduksi insulin yaitu hormon yang berfungsi mengatur gula darah ataupun glukosa yang cukup ketika tubuh tidak memiliki kemampuan untuk menggunakan insulin yang dihasilkan secara efektif. Diabetes termasuk permasalahan kesehatan yang sangat penting bagi masyarakat, karena merupakan satu dari empat penyakit prioritas yang tidak menular tapi menjadi target untuk ditindak lanjuti oleh para pemimpin dunia[5].

### 2.2. Data mining

*Data Mining* adalah suatu teknik untuk menganalisis sekumpulan data guna mengetahui hubungan yang ada lalu meringkas data sehingga mudah dipahami serta memiliki manfaat untuk pemilik data[6].

### 2.3. Algoritma C4.5

Algoritma C4.5 termasuk kelompok algoritma *Decision Tree* dan memiliki input yang berupa *samples* juga *training samples*[7].

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{S} * \text{Entropy}(S) \tag{1}$$

Dimana:

S : ruang sample dalam proses training.

A : atribut.

|S<sub>i</sub>| : jumlah sample dengan nilai V.

|S| : jumlah keseluruhan sampel data.

Entropi(S<sub>i</sub>) : entropi dari sample-sample dengan nilai i

### 2.4. Algoritma Naive Bayes

*Naive Bayes* adalah metode yang tidak mempunyai aturan, *naive bayes* menggunakan cabang matematika yaitu teori probabilitas dengan memanfaatkan peluang terbesar pada kemungkinan klasifikasi, adapun cara menghitung frekuensi disetiap klasifikasi menggunakan data *training*[3].

Teorema *bayes* memiliki bentuk umum seperti dibawah ini :

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \tag{2}$$

Keterangan:

Y : kelas data yang belum diketahui

X : hipotesis data y dari kelas spesifik

P(x|y) : probabilitas hipotesis dari x sesuai kondisi y atau *posteriori probability*

P(x) : probabilitas hipotesis dari x atau *prior probability*

P(y|x) : probabilitas y sesuai kondisi hipotesis x

P(y) : probabilitas y.

### 2.5. Java

Java merupakan bahasa pemrograman yang dikembangkan oleh James Gosling dengan dibantu oleh rekan-rekannya pada tahun 1991 di perusahaan yang bernama Sun Microsystems. Pada awalnya bahasa pemrograman ini dinamai “Oak”, hingga berganti nama menjadi “Java” pada tahun 1995 [8].

### 2.6. NetBeans IDE

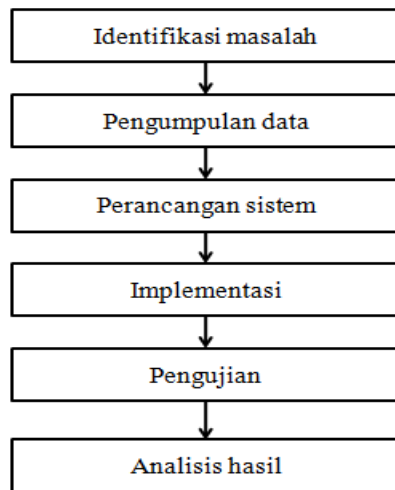
Netbeans merupakan aplikasi *Integrated Development Environment* (IDE) yang berbasisan Java dari Sun Microsystems. NetBeans IDE dimanfaatkan sebagai sebuah alat untuk menulis dan mengompilasi untuk mencari error lalu membagikan program yang ditulis dalam bahasa Java untuk mendukung bahasa pemrograman lainnya.

## 3. METODE PENELITIAN

Metodologi penelitian merupakan suatu proses yang dilakukan secara ilmiah guna mendapatkan data dan hasil penelitian. Metodologi juga melingkupi analisis secara teoretis terhadap metode yang digunakan.

### 3.1. Gambaran umum

Adapun gambar diagram alur dalam penelitian ini dilihat gambar tersebut.

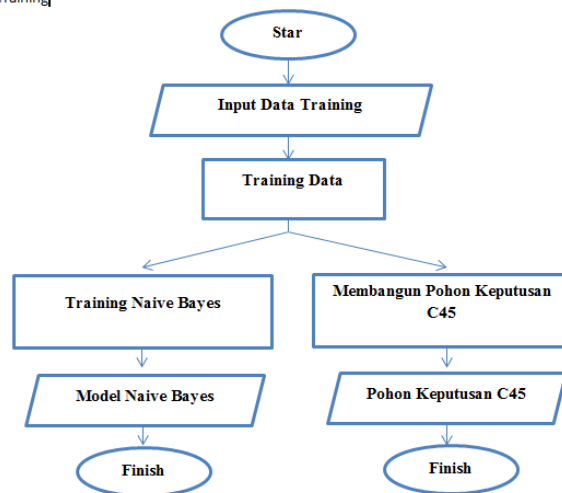


Gambar 1 Bagan langkah penelitian

3.2. Rancangan Sistem

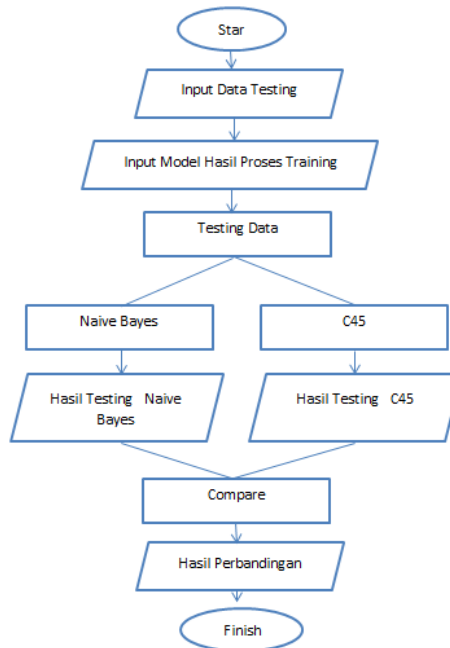
3.2.1. Flowchart Proses Training

Proses Training



Gambar 2 Flowchart Proses Training

3.2.2. *Flowchart Proses Testing*



Gambar 2 *Flowchart* Proses Testing

3. HASIL DAN PEMBAHASAN

Penelitian ini memanfaatkan data dari Puskesmas Lembang Majene tahun 2019 yang ambil secara acak. Data yang diperoleh kemudian dilakukan uji coba sistem dengan menggunakan 2 Algoritma yaitu algoritma NAIVE BAYES dan C45 yang bertujuan untuk melihat perbandingan untuk melahirkan rekomendasi sistem keputusan yang digunakan untuk memperdiksi penyakit Diabetes utamanya di Puskesmas Lembang Majene. Berikut tabel yang digunakan pada penelitian ini.

Tabel 1 Atribut penelitian

Nama Atribut	Penjelasan
Gula darah	Rendah (<95) Sedang (95-140) Tinggi (>140)
Tekanan Darah	Rendah (<100) Sedang (100-140) Tinggi (>140)
Berat Badan	Kurus (<45) Normal (45-50) Gemuk (50-60) Sangat Gemuk(>60)
Usia	Muda (<45) Dewasa (45-59) Tinggi (>60)
Jenis Kelamin	Perempuan (P) Laki-laki (L)
Status (Positif / Negatif)	Positif Negatif

Data atribut digunakan pada saat pengelompokan data pasien penderita penyakit yang diperoleh dari puskesmas Lembang Majene. Pengelompokan data ini dilakukan guna mempermudah pada saat penghitungan algoritma Naive Bayes dan algoritma C45.

3.1. *Data pasien Diabetes*

Penelitian ini menggunakan data pasien penyakit diabetes di puskesmas lembang Majene dengan jumlah data sebanyak 95 record.

Tabel 2 Data Pasien Diabetes

No	Gula Darah	Tekanan Darah	Berat Badan	Umur	Jenis Kelamin	Status
1	112	165	38	86	P	N
2	89	172	43	85	L	N
3	112	168	68	79	P	N
4	440	140	48	51	P	N

Data pasien penyakit diabetes tersebut dikelompokkan berdasarkan atribut dan menjadi seperti ditabel berikut :

Tabel 3 Data pasien yang sudah dikelompokkan

Gula Darah	Tekanan Darah	Berat Badan	Umur	Jenis Kelamin	Status
Sedang	Tinggi	Kurus	Tua	P	N
Rendah	Tinggi	Kurus	Tua	L	N
Sedang	Tinggi	Sangat gemuk	Tua	P	N
Tinggi	Normal	Normal	Dewasa	P	N

3.2. *penghitungan Algoritma Naive Bayes*

Berikut ini merupakan contoh kalkulasi manual Algoritma Naive Bayes dalam mengklasifikasikan penderita penyakit diabetes melitus dengan memanfaatkan data training dan data testing seperti berikut ini:

Menghitung data training

- Menghitung probabilitas variabel Y (peluang status / P(H) )
  - $P(Y = \text{Positif}) = \frac{9}{95} = 0,09473684211$
  - $P(Y = \text{Negatif}) = \frac{86}{95} = 0,90526315789$
- Menghitung probabilitas variabel X terhadap Y / P ( X | H)
  - $P(GD / X_1 \text{ Rendah}) = \text{Positif} = \frac{0}{9} = 0 \rightarrow \frac{1}{10} = 0.1$
  - $\text{Negatif} = \frac{25}{86} = 0.29069767441860467$
  - $\text{Sedang} = \text{Positif} = \frac{1}{9} = 0.1111111111111111$
  - $\text{Negatif} = \frac{46}{86} = 0.5348837209302325$
  - $\text{Tinggi} = \text{Positif} = \frac{8}{9} = 0.8888888888888888$
  - $\text{Negatif} = \frac{15}{86} = 0.1744186046511628$
  - $P(TD / X_2 \text{ Rendah}) = \text{Positif} = \frac{1}{9} = 0.1111111111111111$

$$\begin{aligned}
 & \text{Negatif} = \frac{0}{86} = 0 \rightarrow \frac{1}{87} = 0.01149 \\
 \text{Normal) = Positif} & = \frac{5}{9} = 0.5555555555555556 \\
 & \text{Negatif} = \frac{56}{86} = 0.6511627906976745 \\
 \text{Tinggi) = Positif} & = \frac{3}{9} = 0.3333333333333333 \\
 & \text{Negatif} = \frac{30}{86} = 0.3488372093023256 \\
 \text{P (BB / } X_3 \text{ Rendah ) = Positif} & = \frac{2}{9} = 0.2222222222222222 \\
 & \text{Negatif} = \frac{17}{86} = 0.19767441860465115 \\
 \text{Normal) = Positif} & = \frac{1}{9} = 0.1111111111111111 \\
 & \text{Negatif} = \frac{20}{86} = 0.23255813953488372 \\
 \text{Gemuk) = Positif} & = \frac{4}{9} = 0.4444444444444444 \\
 & \text{Negatif} = \frac{28}{86} = 0.2441860465116279 \\
 \text{Sangat Gemuk) = Positif} & = \frac{2}{9} = 0.2222222222222222 \\
 & \text{Negatif} = \frac{28}{86} = 0.32558139534883723 \\
 \text{P (U / } X_4 \text{ Muda ) = Positif} & = \frac{5}{9} = 0.5555555555555556 \\
 & \text{Negatif} = \frac{17}{86} = 0.19767441860465115 \\
 \text{Dewasa) = Positif} & = \frac{2}{9} = 0.2222222222222222 \\
 & \text{Negatif} = \frac{43}{86} = 0.5 \\
 \text{Tua) = Positif} & = \frac{2}{9} = 0.2222222222222222 \\
 & \text{Negatif} = \frac{26}{86} = 0.3023255813953488 \\
 \text{P (JK / } X_4 \text{ PEREMPUAN ) = Positif} & = \frac{6}{9} = 0.6666666666666666 \\
 & \text{Negatif} = \frac{56}{86} = 0.6511627906976745 \\
 \text{LAKI-LAKI) = Positif} & = \frac{3}{9} = 0.3333333333333333 \\
 & \text{Negatif} = \frac{30}{86} = 0.3488372093023256
 \end{aligned}$$

Tabel 4 Testing Naive Bayes

Gula Darah	Tekanan Darah	Berat Badan	Umur	Jenis Kelamin	Status
125	150	55	59	P	?

Tabel 5 Testing Naive Bayes

Gula Darah	Tekanan Darah	Berat Badan	Umur	Jenis Kelamin	Status
Sedang	Normal	Gemuk	Dewasa	P	?

Proses testing

$$\begin{aligned}
 P ( GD = Sedang ) &= \text{Positif} = 0,11111 \\
 &\text{Negatif} = 0,49473 \\
 P ( TD = Normal ) &= \text{Positif} = 0,04210 \\
 &\text{Negatif} = 0,55789 \\
 P ( BB = Gemuk ) &= \text{Positif} = 0,05263 \\
 &\text{Negatif} = 0,24210 \\
 P ( U = Dewasa ) &= \text{Positif} = 0,02105 \\
 &\text{Negatif} = 0,031557 \\
 P ( JK = P ) &= \text{Positif} = 0,06315 \\
 &\text{Negatif} = 0,58947
 \end{aligned}$$

Format testing

$$\begin{aligned}
 C^* &= \arg \max P ( Y ) \cdot \prod_{i=1}^n P ( X_i | Y ) \\
 C^U / Y &= \text{Positif} \rightarrow P ( Y = \text{positif} ) \cdot \prod_{i=1}^n P ( X_i | Y ) \\
 &= P ( Y . \text{positif} ) \cdot ( P ( X_1 | Y_P ) \cdot P ( X_2 | Y_P ) \dots P ( X_6 | Y_P ) ) \\
 &= 0,09473 ( 0,11111 \times 0,04210 \times 0,05263 \times 0,02105 \times 0,06315 ) \\
 &= 0,000000031 \\
 C^U / Y &= \text{Negatif} \rightarrow P ( Y = \text{Negatif} ) \cdot \prod_{i=1}^n P ( X_i | Y ) \\
 &= P ( Y . \text{Negatif} ) \cdot ( P ( X_1 | Y_P ) \cdot P ( X_2 | Y_P ) \dots P ( X_6 | Y_P ) ) \\
 &= 0,90526 ( 0,49473 \times 0,55789 \times 0,24210 \times 0,031557 \times 0,58947 ) \\
 &= 0,00112523274
 \end{aligned}$$

Jadi, hasil dari proses testing adalah negatif, karena nilai negatif lebih tinggi dari pada nilai positif.

3.3. *penghitungan Algoritma C45*

Perhitungan dengan memanfaatkan algoritma C4.5 memiliki perbedaan dengan algoritma Naive Bayes. Algoritma C4.5 memanfaatkan pohon keputusan untuk rule yang akan digunakan untuk mengklasifikasikan penderita penyakit Diabetes Militus.

1. Perhitungan jumlah kasus. Jumlah kasus bagi keputusan positif, jumlah kasus bagi keputusan negatif, serta entropy bagi semua kasus. Kasus dibagi berdasarkan atribut GD ( gula darah), TD ( tekanan darah ), BB ( berat badan ), usia, jenis kelamin. Setelah itu, perhitungan gain dilakukan untuk setiap atribut dan ditunjukkan oleh tabel berikut :

Tabel 6 Perhitungan C45 Node 1

Node	Jumlah Kasus	Negatif /S1	Positif /S2	Entropi	Gain
Total Gula	95	86	9	0,452086199	
1 Darah					-0,331963538
Tinggi	23	15	8	0,932111568	
Sedang	47	46	1	0,14854949	
Rendah	25	25	0	0	
2 Tekanan Darah					0,644745083
Tinggi	33	30	3	0,197770213	
Normal	60	55	5	0,41381685	
Rendah	2	2	0	0	



3	Berat Badan					0,733000316
		Sangat Gemuk	30	28	2	0,353359335
		Gemuk	25	21	4	0,634309555
		Normal	21	20	1	0,276195428
		Kurus	18	17	1	0,309543429
4	Umur					0,506691706
		Tua	28	26	2	0,371232327
		Dewasa	45	43	2	0,26231122
		Muda	22	17	5	0,773226674
6	Jenis Kelamin					0,305400725
		P	62	56	6	0,458685816
		L	33	30	3	0,439496987

Sesuai tabel diatas, contoh perhitungan entropy serta gain pada field Kgp ( kadar glukosa puasa) dengan mengaplikasikan algoritma C4.5 adalah sebagai berikut:

$$\begin{aligned} \text{Entropy ( Total )} &= \left(-\left(\frac{86}{95} * \log_2 \left(\frac{86}{95}\right)\right) + \left(-\left(\frac{9}{95} * \log_2 \left(\frac{9}{95}\right)\right)\right)\right) \\ &= 0,12998751 + 0,322098689 \\ &= 0,452086199 \end{aligned}$$

$$\begin{aligned} \text{Entropy (High )} &= \left(-\left(\frac{15}{23} * \log_2 \left(\frac{15}{23}\right)\right) + \left(-\left(\frac{8}{23} * \log_2 \left(\frac{8}{23}\right)\right)\right)\right) \\ &= 0,402176974 + 0,529934593 \\ &= 0,932111568 \end{aligned}$$

$$\begin{aligned} \text{Entropy (Medium )} &= \left(-\left(\frac{46}{47} * \log_2 \left(\frac{46}{47}\right)\right) + \left(-\left(\frac{1}{47} * \log_2 \left(\frac{1}{47}\right)\right)\right)\right) \\ &= 0,030366749 + 0,118182742 \\ &= 0,14854949 \end{aligned}$$

$$\begin{aligned} \text{Gain (Gula Darah)} &= 0,452086199 - \left(\left(\frac{23}{95} * 0,932111568\right) + \left(\frac{47}{95} * 0,932111568\right) + \left(\frac{25}{95} * 0\right)\right) \\ &= - 0,405456444 + 0,378593293 + 0,452086199 \\ &= - 0,331963538 \end{aligned}$$

### 3.4. Pengujian Data

Percobaan dilakukan sebanyak dua kali untuk membandingkan algoritma Naive Bayes juga C4.5 untuk memprediksi penderita penyakit Diabetes Melitus pada pasien di Puskesmas Lembang Majene. Dengan menggunakan data testing, diperoleh persentase akurasi yang berbeda-beda. Masing – masing algoritma yang diuji memiliki rata – rata yaitu 74.31% untuk pengujian akurasi algoritma naive bayes dan 90,52% untuk pengujian algoritma C4.5.

## 4. KESIMPULAN

Hasil penelitian yang telah dilakukan oleh peneliti menunjukkan bahwa dari hasil eksperimen dengan membandingkan algoritma Naive Bayes dengan C4.5 dalam melakukan prediksi terhadap penderita penyakit Diabetes Melitus bagi pasien di Puskesmas Lembang Majene menghasilkan rata-rata akurasi 74,31 % untuk algoritma Naive Bayes dan rata-rata

akurasi 90.52% untuk algoritma C4.5 untuk data testing dengan 95 data training. Dari hasil tersebut dapat disimpulkan bahwa algoritma yang paling baik dalam mengklasifikasikan penderita penyakit Diabetes Melitus bagi pasien di Puskesmas Lembang Majene adalah algoritma C4.5.

Adapun peningkatan kinerja sistem serta untuk menyempurnakan penelitian terkait, diharapkan penelitian berikutnya dapat membandingkan kedua algoritma dengan algoritma klasifikasi yang lain untuk memperoleh informasi terkait algoritma mana yang lebih baik dalam melakukan prediksi terhadap penderita penyakit Diabetes Militus khususnya di Puskesmas Lembang Majene. Selain itu, diharapkan penambahan data yang akan digunakan untuk proses training, sehingga data training menjadi lebih banyak lagi sehingga tingkat keakurasiannya dari metode yang digunakan bisa menjadi lebih baik.

#### REFERENSI

- [1] N. A. Nnamoko, F. N. Arshad, D. England, and J. Vora, "Meta-classification model for diabetes onset forecast: A proof of concept," in *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE Computer Society, 2014, pp. 50–56. doi: 10.1109/BIBM.2014.6999247.
- [2] D. Garnita, "Faktor Risiko Diabetes Melitus di Indonesia (Analisis Data Sakerti 2007)," Universitas Indonesia, Depok, 2012.
- [3] S. Bahri, D. M. Midyanti, and R. Hidayati, "Perbandingan Algoritma Naive Bayes dan C4.5 Untuk Klasifikasi Penyakit Anak," in *SNATi 2018: Seminar Nasional Teknik Informatika*, Yogyakarta: Jurusan Teknik Informatika Fakultas Teknologi Industri Universitas Islam Indoensia, 2018, pp. B24–B31.
- [4] H. Dhika, "Belajar Mudah Algoritma Data Mining C4.5," *ilmubiner*, 2014. <http://ilmubiner.blogspot.com/2014/02/belajar-mudah-algoritma-data-mining-c45.html>
- [5] R. N. Fatimah, "Diabetes Melitus Tipe 2," vol. 4, no. 5 (2015), pp. 93–101, Feb. 2015.
- [6] Kusri and E. T. Luthfi, *Algoritma Data Mining*. Yogyakarta: Penerbit Andi, 2009. [Online]. Available: [http://ucs.sulsellib.net//index.php?p=show\\_detail&id=151247](http://ucs.sulsellib.net//index.php?p=show_detail&id=151247)
- [7] A. Saikhu, J. Lianto, and U. Hanik, "Fuzzy Decision Tree dengan Algoritma C4.5 pada Data Diabetes Indian Pima," presented at the Konferensi Nasional Sistem dan Informatika, Bali, Bali, 2011, pp. 297–302. [Online]. Available: <https://yudiagusta.files.wordpress.com/2008/09/297-302-knsi2011-047-fuzzy-decision-tree-dengan-algoritma-c4-5-pada-data-diabetes-indian-pima.pdf>
- [8] A. Nugroho, *Pemrograman Java Menggunakan IDE Eclipse Callisto*. 2008.