

KLASIFIKASI DIAGNOSIS DIABETES MELITUS MENGUNAKAN METODE NAÏVE BAYES DENGAN SELEKSI FITUR *BACKWARD ELIMINATION*

DIABETES MELITUS DIAGNOSIS CLASSIFICATION USING THE NAIVE BAYES METHOD WITH FEATURE SELECTION BACKWARD ELIMINATION

Hendro Nugroho¹⁾, Gusti Eka Yuliasuti²⁾, Andrian Firman P³⁾

^{1), 2), 3)}Program Studi Teknik Informatika, Fakultas Teknik Elektro dan Teknologi Informasi, Institut
Teknologi Adhi Tama Surabaya

Jl Arief Rahman Hakim no 100 Surabaya

Email : dosh3ndro@itats.ac.id¹⁾, gustiekay@itats.ac.id²⁾

Abstrak

Diabetes melitus merupakan penyakit yang berbahaya disebabkan dari kadar gula yang tinggi (hiperglikemia). Hiperglikemia yang dapat menyebabkan penderita mengalami penyakit kronis kerusakan pada organ pada tubuh. Diabetes melitus yang menjadi penyakit berbahaya, maka sangatlah menarik melakukan klasifikasi Diabetes melitus menggunakan metode Naïve bayes dengan seleksi fitur Backward Elimination (BE). Dataset Diabetes melitus yang digunakan penelitian sejumlah 101 data dengan 5 atribut yang terdiri dari umur, Gula Darah Sewaktu (GDS), 2 jam setelah makan/Post Pradial (PP), kadar Gula Darah Puasa (GPD), dan Low Density Lipoprotein (LDL). Untuk mendapatkan hasil klasifikasi terdapat beberapa langkah-langkah yang dilakukan yaitu input data, seleksi fitur BE, 8-Fold Cross Validation, Naïve bayes dan pengujian hasil. Dari hasil klasifikasi dilakukan pengujian menggunakan metode penghitungan akurasi, precision dan recall. Untuk mengetahui hasil kinerja klasifikasi dilakukan empat skenario pengujian yaitu skenario pertama Naïve bayes dikombinasikan dengan BE dan 8-Fold Cross Validation akurasi sebesar 77%, skenario kedua Naïve bayes dikombinasikan dengan 8-Fold Cross Validation akurasi sebesar 78.1%, skenario ketiga Naïve Bayes dikombinasikan dengan BE akurasi sebesar 86% dan skenario keempat klasifikasi Naïve bayes akurasi sebesar 90%, sehingga akurasi klasifikasi Naïve Bayes dengan seleksi fitur BE menjadi lebih baik.

Kata kunci: Diabetes melitus, Naive Bayes, Backward Elimination, 8-Fold Cross Validation

Abstract

Diabetes mellitus is a dangerous disease caused by high sugar levels (hyperglycemia). Hyperglycemia can cause sufferers to experience chronic disease, damage to organs in the body. Diabetes mellitus is a dangerous disease, so it is very interesting to classify diabetes mellitus using the Naïve Bayes method with Backward Elimination (BE) feature selection. The Diabetes mellitus dataset used in the research consisted of 101 data with 5 attributes consisting of age, Current Blood Sugar (GDS), 2 hours after eating/Post Pradial (PP), Fasting Blood Sugar (GPD) levels, and Low Density Lipoprotein (LDL). To get classification results, there are several steps taken, namely data input, BE feature selection, 8-Fold Cross Validation, Naïve Bayes and results testing. From the classification results, testing was carried out using the accuracy, precision and recall calculation method. To find out the results of classification performance, four test scenarios were carried out, namely the first scenario, Naïve Bayes combined with BE and 8-Fold Cross Validation, accuracy of 77%, second scenario, Naïve Bayes combined with 8-Fold Cross Validation, accuracy of 78.1%, third scenario, Naïve Bayes combined with BE accuracy is 86% and the fourth scenario of Naïve Bayes classification accuracy is 90%, so the accuracy of Naïve Bayes classification with BE feature selection is better.

Keywords: Diabetes melitus, Naïve bayes, Backward Elimination. 8-Fold Cross Validation.

1. PENDAHULUAN

Diabetes melitus merupakan penyakit yang berbahaya yang mengancam kesehatan manusia dan penderita diabetes melitus mengalami peningkatan yang tinggi di seluruh dunia[1] dan yang paling parah melanda negara-negara berpenghasilan rendah dan menengah[2]. Penderita Diabetes melitus ditandai oleh hiperglikemia[3]. Hiperglikemia dapat menyebabkan penderita mengalami penyakit kronis kerusakan berbagai organ dan disfungsi jaringan yang dapat

menyebabkan hipertensi, kardiovaskuler, gagal ginjal, dan komplikasi lainnya[1]. Berbahayanya penyakit Diabetes melitus yang menjadi salah satu prioritas penyakit tidak menular yang harus diperhatikan dan ditindaklanjuti oleh pemerintah seluruh dunia[4].

Penyakit diabetes melitus yang sangat berbahaya inilah yang mendorong peneliti melakukan penelitian. Penelitian sebelumnya menggunakan data diabetes melitus dengan menggunakan algoritma Naïve bayes menghasilkan klasifikasi dataset Diabetes dengan akurasi 90,20%[5]. Pada penelitian lainnya membandingkan algoritma Naïve bayes dan K-Nearest Neighbor (KNN) pada data penyakit diabetes, hasil dari penelitian tersebut mendapatkan hasil akurasi klasifikasi sebesar 85.60% untuk Naïve bayes dan 91.61% untuk KNN[6].

Dari penelitian sebelumnya penggunaan *machine learning* merupakan hal yang menarik untuk dilakukan penelitian berikutnya dengan menggunakan data Diabetes melitus. Didalam penelitian ini bertujuan untuk memperbaiki kelemahan yang dihadapi di dalam klasifikasi Naïve bayes. Di dalam penelitian sebelumnya untuk perbaikan klasifikasi dari data kesehatan dengan seleksi fitur untuk mengatasi atribut yang tidak sesuai dengan menggunakan metode *elimination* fitur[7]. Penggunaan klasifikasi Naïve bayes dengan menggunakan metode seleksi fitur dapat memperbaiki kinerja klasifikasi dengan baik[8]. Pada penelitian lain menggunakan metode *elimination* seleksi fitur pada data kesehatan fitur hidung elektronik (E-Nose) menggunakan metode *Backward Elimination* dapat meningkatkan hasil akurasi klasifikasi[9].

Pada permasalahan diatas, maka penelitian ini melakukan klasifikasi Diabetes Melitus menggunakan metode Naïve bayes dengan seleksi fitur menggunakan metode *Backward Elimination*. Penggunaan metode tersebut sebagai metode optimasi untuk meningkatkan kinerja klasifikasi.

2. DASAR TEORI

Penelitian klasifikasi Diabetes melitus menggunakan metode Naïve Bayes dilandasi dengan berbagai teori. Dasar teori yang digunakan dalam penelitian ini meliputi dataset Diabetes melitus, Naïve bayes, K-Fold Cross validasi, Seleksi fitur, dan pengujian.

2.1. Diabetes Melitus

Diabetes melitus atau dikenal dengan kencing manis adalah penyakit yang disebabkan oleh kadar gula darah yang tinggi (hiperglikemia)[10] dan berbagai macam gejala penyakit lainnya yang berlangsung dalam jangka waktu lama [11]. Diabetes melitus menjadi penyakit yang berbahaya di dunia yang perlu dilakukan penanganan serius. Penyebab Diabetes melitus dari kebiasaan makan yang tidak sehat, kurang olahraga, penuaan, riwayat keluarga, stres, kecemasan dan depresi[11].

Pengukuran penyakit Diabetes melitus bisa dijadikan variabel untuk dataset penelitian. Pengukuran gejala penyakit Diabetes melitus dapat diambil berdasarkan umur(penuaan), kadar gula selama sehari (GDS), kadar gulah darah yang diambil 2 jam setelah makan, (2 jampp), kadar gula setelah puasa (GPD) dan kadar kolesterol (LDL). Data yang dihasilkan untuk pengukuran Diabetes melitus bersifat numerik, sehingga dapat digunakan untuk menentukan pengukuran dengan menggunakan metode data mining. Data pengukuran tersebut dapat digunakan sebagai data *training* dan data uji.

2.2. Metode Naïve Bayes

Metode Naïve bayes merupakan metode *Machine learning* yang digunakan untuk klasifikasi pada data set. Penggunaan metode Naïve bayes merupakan metode dengan penggunaan *learning* yang tercepat dan mudah[12]. Naïve bayes merupakan klasifikasi yang menggunakan cara probabilistik simpel dengan penjumlahan frekuensi serta nilai dari dataset yang ada.[13]. Probabilitas pada teorema bayesian, mengasumsikan setiap variabel memiliki sifat independen, dengan kata lain keberadaan suatu atribut tidak memiliki kesamaan dengan atribut lain[14].

Untuk mendapatkan nilai probabilitas Naïve bayes menggunakan persamaan 1.

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \dots\dots(1)$$

Dimana,

X : data *class* yang belum diketahui

H : Hipotesis *class* spesifik

P(H|X) : probabilitas hipotesis H berdasarkan kondisi X

P(X|H) : probabilitas X berdasarkan hipotesis H

P(X) : probabilitas X

2.3. Cross Validation

Penggunaan *Cross-validation* digunakan untuk generalisasi model *learning*[15]. Untuk mendapatkan model yang dijadikan data statistik menggunakan metode *K-Fold Cross Validation*. Metode *K-Fold Cross Validation* membagi dua data menjadi data uji dan data training. Data training untuk membentuk melatih model klasifikasi dan data uji dibentuk untuk menguji model klasifikasi[16].

Kinerja dari *K-Fold Cross Validation* menjalankan nilai k pada interasi pada data. K merupakan jumlah node untuk dilatih. Nilai-nilai k yang digunakan k=5 atau k=10 untuk *K-Fold Cross Validation*. Nilai-nilai k yang digunakan secara empiris menghasilkan perkiraan kesalahan yang tidak terlalu tinggi. Dalam menentukan nilai k dapat menghasilkan nilai pengujian secara baik.

2.4. Seleksi Fitur Backward Elimination

Seleksi fitur digunakan untuk proses *machine learning* yang memiliki atribut yang banyak. Tujuan dari seleksi fitur adalah mengurangi data yang tidak digunakan, sehingga dapat mempercepat komputasi. Seleksi fitur membuat klasifikasi lebih efisien dan efektif dengan mengurangi jumlah data yang sesuai dalam proses *learning*[17], berfungsi sebagai seleksi atribut dimana memanfaatkan regresi statistik untuk mengetahui kedekatan setiap kombinasi atribut dengan target.. Dataset Diabetes melitus memiliki atribut yang banyak, maka dilakukan proses seleksi fitur. Untuk metode seleksi fitur pada penelitian ini menggunakan metode *Backward Elimination*.

Backward Elimination (BE) adalah algoritma yang digunakan untuk memilih fitur dari serangkaian kombinasi fitur untuk menemukan kombinasi terbaik secara rekursif. Cara kerja algoritma BE dengan menguji semua fitur dan kemudian secara bertahap mengurangi fitur yang tidak signifikan[9]. Nilai signifikan level untuk pemilihan atribut yang akan dipilih. Semakin kecil nilai signifikan level semakin ketat pemilihan atribut yang akan terpilih, maka semakin sedikit atribut sebagai model[17].

2.5. Pengujian Hasil Klasifikasi

Pengujian hasil klasifikasi Diabetes melitus menggunakan metode Naïve Bayes untuk mengetahui nilai akurasi dari penelitian ini. Metode yang digunakan untuk mendapatkan nilai akurasi menggunakan metode *Confusion Matrix* (CM). Metode CM terdiri dari kelas *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN). Dari kelas-kelas tersebut mendapatkan nilai akurasi, *precision*, dan *recall*.

Nilai akurasi adalah nilai efektivitas dari kinerja model klasifikasi secara keseluruhan[7]. Nilai akurasi menunjukkan rasio klasifikasi yang benar dari total pengamatan data penelitian. Untuk mendapatkan nilai akurasi dapat dilihat pada persamaan 2.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots(2)$$

Nilai *precision* untuk nilai ketepatan kinerja model klasifikasi dan untuk menyatakan klasifikasi nilai positif atas prediksi nilai positif[7]. Perhitungan nilai *precision* dapat dilihat pada persamaan 3.

$$Precision = \frac{TP}{TP+FP} \dots\dots(3)$$

Nilai *recall* adalah nilai rata-rata *true positive* untuk menghitung seberapa baik model dalam mengidentifikasi class normal[7]. Untuk mendapatkan nilai *recall* dapat dilihat pada persamaan 4.

$$Recall = \frac{TP}{FN+TP} \dots\dots(4)$$

Dimana,

TP adalah nilai jumlah data positif yang telah diklasifikasikan sebagai positif

FP adalah nilai jumlah data negatif yang telah diklasifikasikan sebagai positif

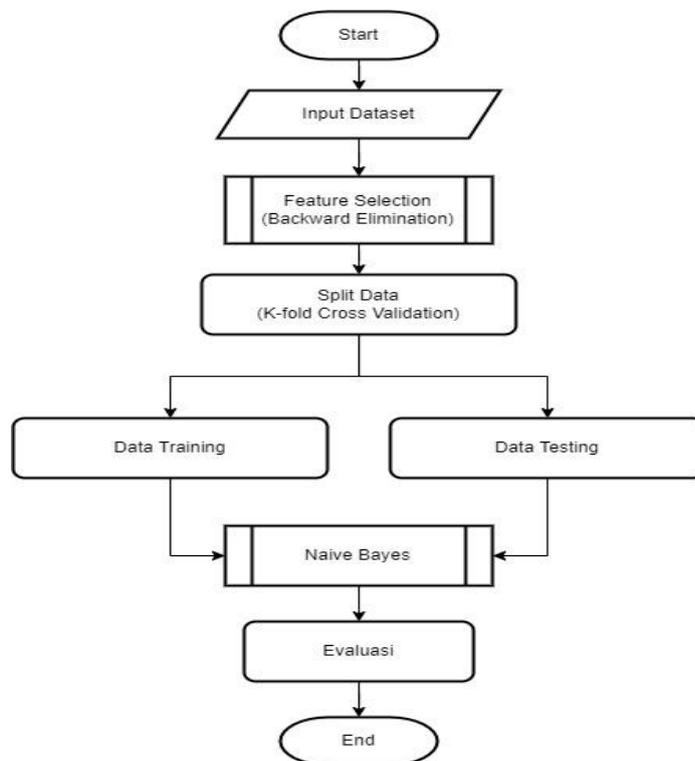
FN adalah nilai jumlah data positif yang telah diklasifikasikan sebagai negatif.

TN adalah nilai jumlah data negatif yang telah diklasifikasikan sebagai negatif.

3. METODOLOGI PENELITIAN

Metode Naive bayes digunakan untuk klasifikasi Diabetes melitus. Untuk mendapatkan hasil klasifikasi pada penelitian ini dilakukan tahapan-tahapan penelitian yang digambarkan dalam bentuk flowchart. Tahapan yang dilakukan di dalam penelitian ini adalah pengumpulan data Diabetes melitus, fitur seleksi menggunakan metode *Backward Elimination*, *cross* validasi menggunakan metode *k-fold cross validation*, klasifikasi Naive bayes, dan pengujian hasil menggunakan *confusion* matrik.

Flowchart klasifikasi Diabetes melitus menggunakan metode Naive bayes dapat dilihat pada Gambar 1. Pada *flowchart system* klasifikasi Diabetes melitus dapat dijelaskan pada subbab dibawah ini.



Gambar 1. Flowchart System Klasifikasi Naive Bayes

3.1. Pengumpulan Dataset

Dataset di dalam penelitian ini adalah data primer Diabetes melitus sejumlah 101 kasus Diabetes melitus. Pada dataset yang diambil terdapat lima variabel yang terdiri dari umur, GDS, 2 jampp, GPD, LDL dan target biner Diabetes (1 dan 0) di tentukan bersama pakar yaitu dokter. Dasar penentuan target sebagai data output.

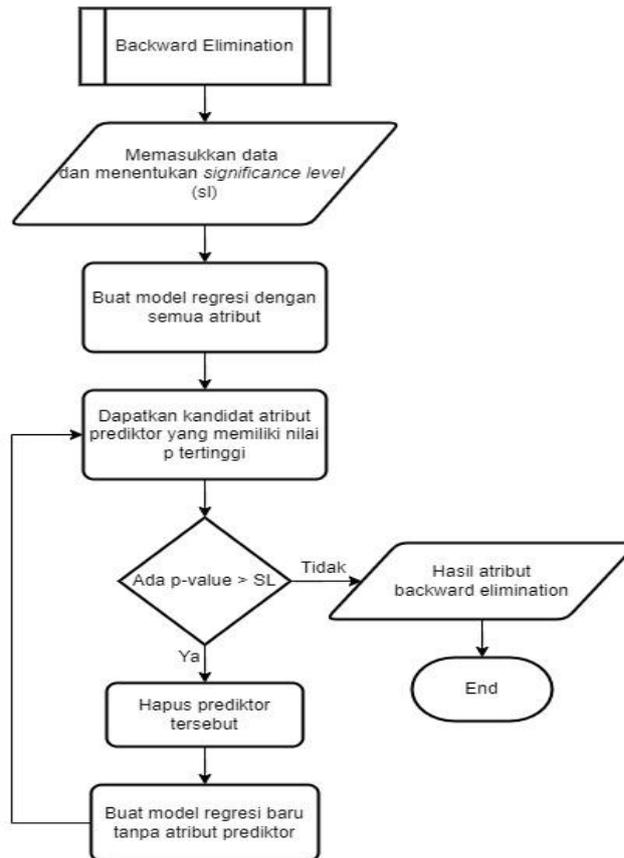
Nilai target biner untuk nilai 1 untuk menyatakan “ya” dan nilai 0 untuk menyatakan “tidak”. Untuk mengetahui contoh dataset Diabetes melitus dapat dilihat pada Tabel 1.

Tabel 1. Contoh Dataset Diabetes Melitus

Age	GDS	2JamPP	GDP	LDL	DM
71	155	222	140	163	1
39	75	108	145	111	1
50	222	222	114	138	1
38	202	215	278	161	1
61	211	203	93	146	1
63	122	111	100	158	0
65	111	172	114	146	0
69	134	160	111	127	0
63	99	126	132	158	0
46	134	123	101	226	0

3.2. Seleksi Fitur Backward Elimination

Pada tahapan seleksi fitur menggunakan metode *Backward Elimination*. Penggunaan metode *Backward Elimination* adalah untuk pemilihan variabel yang menggunakan regresi statistik untuk mengetahui kedekatan setiap kombinasi variabel target. Tahapan metode *Backward Elimination* digambarkan dalam bentuk *flowchart* pada gambar 2.



Gambar 2. Flowchart Seleksi Fitur Backward Elimination

Pada Gambar 2 dapat dijelaskan langkah-langkah seleksi fitur *Backward Elimination* dimulai dari menginputkan dataset dengan variabel umur, GDS, 2 jampp, GDP, LDL dan menentukan signifikan level. Nilai signifikan level dapat menggunakan nilai 0.05 dan 0.1 [17].

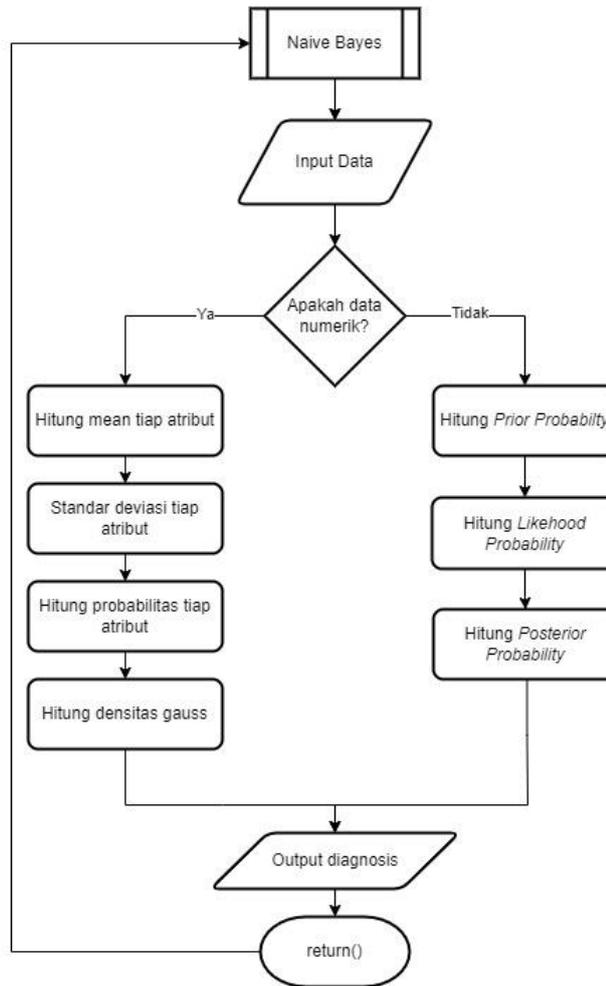
3.3. Split Data (K-Fold Cross Validation)

Penggunaan metode *Cross Validation* untuk membagi dataset menjadi data *training* dan data *testing*. Pembagian data *training* dan data *testing* menggunakan *8-Fold Cross Validation*. Data training dibentuk untuk melatih model klasifikasi dan data uji dibentuk untuk menguji model klasifikasi. *K-Fold Cross Validation* merupakan nilai k iterasi pada data, dimana nilai k merupakan jumlah node yang akan dilatih.

3.4. Metode Naïve Bayes

Setelah dilakukan split data, maka dilakukan proses klasifikasi Naïve bayes. Metode klasifikasi Naïve bayes menggunakan persamaan 1. Teori Naïve bayes merupakan metode klasifikasi berdasarkan probabilitas di mana keberadaan suatu atribut memiliki perbedaan dengan atribut lainnya [14]. Cara kerja metode Naïve bayes dapat digambarkan dalam bentuk *flowchart* pada Gambar 3.

Pada Gambar 3 flowchart metode Naïve bayes dapat dijelaskan alur proses dimulai menginputkan data Diabetes melitus. Langkah selanjutnya pengujian dataset untuk menghitung nilai probabilitas prior yang memiliki dua kelas, yaitu kelas Diabetes dan tidak Diabetes.



Gambar 3. Flowchart Naïve Bayes

3.5. Pengujian Hasil

Pengujian hasil klasifikasi Diabetes melitus bertujuan untuk mengukur tingkat keberhasilan didalam penelitian ini. Nilai pengujian hasil juga dapat digunakan untuk mendapatkan hasil analisa didalam penelitian klasifikasi Diabetes melitus menggunakan metode Naïve Bayes dan seleksi fitur.

Analisa dari hasil penelitian klasifikasi Diabetes melitus didapat dari nilai penghitungan akurasi, *precision*, dan *recall*. Untuk mendapatkan nilai hasil pengujian menggunakan persamaan 2 untuk nilai akurasi, persamaan 3 untuk *precision*, dan persamaan 4 untuk *recall*.

4. PENGUJIAN DAN PEMBAHASAN

Sebelum melakukan pengujian pada penelitian klasifikasi Diabetes melitus menggunakan metode Naïve bayes terdapat empat tahapan. Tahapan pertama adalah pengumpulan data Diabetes melitus, tahap kedua pengolahan data, tahap ketiga pemodelan data dan tahap keempat pengujian model.

4.1. Pengumpulan Data Diabetes Melitus

Data training diperoleh dari data primer Rumah sakit Haji Surabaya tahun 2022 sejumlah 101 data Diabetes melitus. Data Diabetes melitus yang diambil memiliki lima variabel. Jumlah data *training* adalah 101 yang terdiri atas 81 dengan pasien Diabetes melitus dan 20 pasien yang tidak Diabetes melitus. Dari jumlah data 101 dibagi menjadi 80% data *training* dan 20% data *testing*.

Setelah dilakukan pengumpulan data Diabetes melitus dilakukan pengolahan data. metode pengolahan data menggunakan metode seleksi fitur *Backward Elimination* dan uji validasi data.

4.2. Seleksi Fitur dan Validasi

Proses pengolahan data yang menggunakan seleksi fitur *Backward Elimination* yang memanfaatkan regresi statistik untuk mengetahui kedekatan kombinasi atribut yang ditargetkan. Pada penelitian menggunakan metode klasifikasi Naïve Bayes dengan menggunakan seleksi fitur *Backward Elimination* dengan menerapkan *Ordinary Least Square* menyisakan tiga variabel dari 5 variabel. Hasil dari seleksi fitur dapat dilihat pada Tabel 2.

Tabel 2. Hasil *Backward Elimination*

Attributes	Standard Error	t	P> t
Age	0,001	3,082	0,003
GDS	0,001	3,367	0,001
GDP	0,001	2,130	0,036

Hasil pada Tabel 2 terjadi reduksi fitur dari data awal 5 atribut menjadi 3 atribut. Selanjutnya dilakukan pengujian *Cross Validation*. *Cross Validation* menggunakan K-Fold Corss Validation dengan nilai K=1 sampai dengan K=5[18]. Penelitian ini menggunakan nilai K=1 hingga K=8. *Cross Validation* digunakan untuk menentukan akurasi yang memiliki nilai terbaik. Hasil *Cross Validation* dapat dilihat pada Tabel 3.

Tabel 3. Hasil *8-Fold Cross Validation*

Iterasi	Naïve Bayes Classification Accuracy
K1	0.916
K2	1.0
K3	0.75
K4	0.75
K5	0.75
K6	0.5
K7	0.583
K8	0.916
Average	0.771

4.3. Hasil Analisa Klasifikasi

Hasil klasifikasi Diabetes melitus menggunakan metode Naïve bayes dilakukan analisis guna mendapatkan hasil evaluasi. Hasil evaluasi didapat dari nilai akurasi, *precision*, dan *recall*. Hasil penelitian klasifikasi Diabetes melitus menggunakan metode Naïve bayes dengan menggunakan seleksi fitur *Backward Elimination* (BE) dan 8-Fold Validation dengan data split

80/20 menghasilkan akurasi 77%, nilai *precision* sebesar 87,5% dan nilai *recall* 63%. Dari hasil tersebut, maka dilakukan pengujian bentuk model lain untuk mendapatkan hasil evaluasi pada penelitian klasifikasi Diabetes melitus.

Pengujian dalam bentuk model lain berupa skenario pengujian. Skenario pengujian pertama dilakukan klasifikasi hanya menggunakan metode Naïve bayes, skenario kedua metode Naïve bayes menggunakan seleksi fitur BE, skenario ketiga metode Naïve bayes menggunakan *8-Fold Cross Validation*. Hasil dari semua skenario pengujian dapat dilihat pada Tabel 4.

Tabel 4. *Tabel Hasil Pengujian*

<i>Validation Test</i>	<i>NB (Split 80/20)</i>	<i>NB+BE (Split 80/20)</i>	<i>NB (8fold validation)</i>	<i>NB+BE (8fold validation)</i>
<i>Precision</i>	100%	100%	87,5%	87,5%
<i>Recall</i>	88%	82%	66,1%	63%
<i>Accuracy</i>	90%	86%	78,1%	77%

Pada tabel 4 menunjukkan bahwa pada penelitian klasifikasi menggunakan metod Naïve Bayes menggunakan seleksi fitur BE dan *8-Fold Cross Validation* tidak dapat meningkatkan akurasi, karena atribut pada dataset Diabetes melitus dari 5 atribut merupakan atribut yang sangat penting untuk mengklasifikasikan penderita Diabetes melitus. hasil tersebut menunjukkan klasifikasi Naïve bayes mendapatkan nilai akurasi 90%, nilai *precision* 100% dan nilai *recall* sebesar 88%.

5. KESIMPULAN

Penelitian klasifikasi menggunakan dataset Diabetes melitus menggunakan 101 data dengan 5 atribut menghasilkan nilai akurasi, *precision* dan *recall*. Dari hasil pengujian mendapatkan analisis perbandingan, algoritma Naïve Bayes menghasilkan kinerja yang lebih baik dengan nilai akurasi 90% dibandingkan dengan metode Naïve bayes dikombinasikan dengan seleksi fitur *Backward Elimination* sebesar 86%. Dan kombinasi metode dengan menggunakan metode *8-Fold Cross Validation* juga tidak ada peningkatan untuk nilai akurasi sebesar 78,1%.

Hasil didapat dari penelitian klasifikasi Diabetes melitus dengan menggunakan metode Naïve Bayes dikombinasikan metode BE dan *8-Fold Cross Validation* menunjukkan kinerja kurang baik, karena atribut pada dataset cukup kecil (5 atribut). Sehingga pada penelitian selanjutnya bisa digunakan untuk atribut yang cukup banyak.

DAFTAR PUSTAKA

[1] Z. Xu and Z. Wang, “A Risk Prediction Model for Type 2 Diabetes Based on Weighted Feature Selection of Random Forest and XGBoost Ensemble Classifier,” *2019 Elev. Int. Conf. Adv. Comput. Intell.*, pp. 278–283.

[2] P. M. Maina, M. Pienaar, and M. Reid, “International Journal of Nursing Studies Advances Self-management practices for preventing complications of type II diabetes mellitus in low and middle-income countries : A scoping review,” *Int. J. Nurs. Stud. Adv.*, vol. 5, no. May, p. 100136, 2023.

[3] R. A. Alassaf, K. A. Alsulaim, N. Y. Alroomi, N. S. Alsharif, M. F. Aljubeir, and O. Sunday, “Preemptive Diagnosis of Diabetes Mellitus Using Machine Learning,” *2018 21st Saudi Comput. Soc. Natl. Comput. Conf.*, pp. 1–5, 2018.

[4] Khairani, “Hari Diabetes Sedunia Tahun 2018,” in *Pus. Data dan Inf.Kementrian kesehatan*, 2019, pp. 1 – 8.

[5] A. Ridwan, “Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Penyakit Diabetes Mellitus,” vol. IV, no. September, pp. 15–21, 2020.

[6] M. D. Nurmalasari, “Komparasi Algoritma Naïve Bayes dan K-Nearest Neighbor untuk Membangun Pengetahuan Diagnosa Penyakit Diabetes,” vol. 5, no. 1, pp. 52–59, 2021.

-
- [7] S. Islam and M. M. Rahman, "Feature Selection and Classification of Spinal Abnormalities to Detect Low Back Pain Disorder using Machine Learning Approaches," *2019 1st Int. Conf. Adv. Sci. Eng. Robot. Technol.*, vol. 2019, no. Icasert, pp. 1–4, 2019.
- [8] O. Tejaswini, A. S. K. P., K. R. Geethika, and G. R. Brindha, "New Feature Selection Process to Enhance Naïve Bayes Classification," *2018 Second Int. Conf. Electron. Commun. Aerosp. Technol.*, no. ICECA, pp. 98–101, 2018.
- [9] M. Tharmakulasingham, C. Topal, A. Fernando, and R. La Ragione, "Backward Feature Elimination for Accurate Pathogen Recognition Using Portable Electronic Nose," *2020 IEEE Int. Conf. Consum. Electron.*, pp. 1–5, 2020.
- [10] E. A. P. S. Siregar, "Efektivitas perawatan ulkus diabetikum terhadap penerimaan diri pasien diabetes melitus tipe II," *Jumantik*, vol. 4, no. 2, pp. 178 – 187, 2019.
- [11] O. Adeleke, H. Sokolayam, D. Emmanuel, A. Daniel, and A. Busola, "Diabetes mellitus : From molecular mechanism to pathophysiology and pharmacology," *Med. Nov. Technol. Devices*, vol. 19, no. February, p. 100247, 2023.
- [12] F. Liantoni and H. Nugroho, "Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan Knearest Neighbor," *J. Simantec*, vol. 5, no. 1, pp. 9–16, 2015.
- [13] S. Alim, "Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor , Decision Tree Serta Naive Bayes Orange Data Mining Implementation For Student Graduation Classification Using K-Nearest Neighbor ," *Jurnal Ilmiah Nero*, vol. 6, no. 2, pp. 133–144, 2021.
- [14] G. E. Yuliatuti, C. N. Prabiantissa, and A. M. Rizki, "Klasifikasi Penyakit Menular Seksual Menggunakan Naïve Bayes," *INTEGER J. Inf. Technol.*, vol. 7, no. 1, pp. 48–53, 2022.
- [15] A. A. Mahran, R. K. Hapsari, and H. Nugroho, "Penerapan Naive Bayes Gaussian Pada Klasifikasi Jenis Jamur Berdasarkan Ciri Statistik Orde," vol. 5, no. 2, 2020.
- [16] L. R. T, "Penerapan Particle Swarm Optimization (PSO) untuk seleksi atribut dalam meningkatkan akurasi Prediksi Diagnosis Penyakit Hepatitis dengan Metode Algoritma C4.5," *Swabumi*, vol. 4, no. 1, pp. 1–15, 2016.
- [17] I. W. Gamadarenda, I. Waspada, "Implementasi Data Mining Untuk Deteksi Penyakit Ginjal Kronis (PGK) Menggunakan K-Nearest Neighbor (KNN) Dengan Backward Data Mining Implementation For Detection Of Chronic Kidney (CKD) Using K-Nearest Neighbor (KNN) With Backward Elimination," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)* vol. 7, no. 2, pp. 417–426, 2020.
- [18] L. S. Ramdhani, "Penerapan Particle Swarm Optimization (Pso) Untuk Seleksi Atribut Dalam Meningkatkan Akurasi Prediksi Diagnosis Penyakit Hepatitis Dengan Metode Algoritma C4 . 5," *Swabumi*, vol. IV, no. 1, pp. 1–15, 2016.