Indonesian Journal of Multidisciplinary on Social and Technology
Homepage: https://journal.ilmudata.co.id/index.php/ijmst
Vol. 4 No. 3 (2026) pp: 10006-10015
P-ISSN: 2986-6790, e-ISSN: 2986-6782

Analisis Kinerja Algoritma Decision Tree dan Naive Bayes untuk Prediksi Risiko Penyakit Diabetes Menggunakan Rapidminer

Najdah Ibtisamah Shafira Alifah, Haryanto, Harun Al Rasyid
Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika
najdahh.isa@gmail.com*, haryanto.hyt@bsi.ac.id, harun.har@bsi.ac.id

Abstrak

Diabetes melitus merupakan penyakit kronis dengan prevalensi global yang terus meningkat sehingga deteksi dini berbasis data sangat diperlukan. Penelitian ini bertujuan membandingkan kinerja algoritma Decision Tree dan Naive Bayes dalam memprediksi risiko diabetes pada Pima Indians Diabetes Database yang terdiri atas 768 data dengan delapan atribut prediktor dan satu label. Penelitian mengikuti kerangka CRISP-DM dan diimplementasikan menggunakan RapidMiner. Perbedaan dengan penelitian terdahulu terletak pada penggunaan kerangka CRISP-DM, cross validation, serta evaluasi lima metrik sekaligus, termasuk AUC. Tahap persiapan data meliputi pengubahan nilai nol yang tidak valid secara biologis menjadi nilai hilang, imputasi dengan rata-rata atribut, dan penetapan label kelas. Evaluasi menggunakan confusion matrix serta metrik akurasi, precision, recall, F1-score, dan Area Under the Curve (AUC). Hasil menunjukkan Decision Tree memperoleh akurasi 73,57%, precision 68,16%, recall 45,52%, F1-score 54,59%, dan AUC 0,674, sedangkan Naive Bayes memperoleh akurasi 74,48%, precision 64,52%, recall 59,70%, F1-score 62,02%, dan AUC 0,817. Naive Bayes unggul pada empat dari lima metrik, terutama recall dan AUC, yang penting dalam skrining medis karena kegagalan mendeteksi pasien diabetes berdampak serius. Decision Tree tetap unggul pada precision dan menghasilkan model yang mudah ditafsirkan. Temuan ini menunjukkan bahwa Naive Bayes lebih sesuai sebagai dasar sistem pendukung keputusan skrining awal risiko diabetes pada dataset yang digunakan.

Kata kunci: *Naive Bayes, Decision Tree, Prediksi Diabetes, RapidMiner*

Abstract

Diabetes mellitus is a chronic disease whose global prevalence keeps rising, so data-driven early detection is needed. This study compares the performance of the Decision Tree and Naive Bayes algorithms in predicting diabetes risk on the Pima Indians Diabetes Database, which contains 768 records with eight predictor attributes and one label. The study follows the CRISP-DM framework and is implemented in RapidMiner. It differs from earlier studies by combining CRISP-DM, 10-fold cross validation, and five evaluation metrics at once, including AUC. Data preparation converted biologically invalid zero values into missing values, imputed them with attribute means, and set the class label. Evaluation used a confusion matrix with accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). Decision Tree achieved accuracy of 73.57%, precision of 68.16%, recall of 45.52%, F1-score of 54.59%, and AUC of 0.674, while Naive Bayes achieved accuracy of 74.48%, precision of 64.52%, recall of 59.70%, F1-score of 62.02%, and AUC of 0.817. Naive Bayes was better on four of five metrics, especially recall and AUC, which matter in medical screening because missing a diabetic patient has serious consequences. Decision Tree remained better on precision and produced an easily interpreted model. These findings indicate that Naive Bayes is more suitable as the basis of a decision support system for early diabetes risk screening on this dataset.

Keywords : Naive Bayes, Decision Tree, Prediksi Diabetes, RapidMiner

1. Pendahuluan

Perubahan pola hidup akibat kemajuan teknologi, urbanisasi, dan dinamika sosial melahirkan gaya hidup modern yang minim gerak. Konsumsi makanan tinggi lemak, minuman bergula, dan garam berlebih, ditambah kebiasaan merokok dan konsumsi alkohol, dapat memicu resistensi insulin dan gangguan pankreas yang berujung pada diabetes melitus (Nuryamin & Risyda, 2025). Laporan International Diabetes Union tahun 2021 mencatat sekitar 537 juta penderita diabetes di dunia, dan jumlah itu diperkirakan mencapai 783 juta orang pada tahun 2045 (Depari et al., 2025). Karena penyakit ini kompleks dan prevalensinya naik setiap tahun, deteksi dini menjadi kunci untuk menekan risiko komplikasi jangka panjang (Febrian et al., 2024).

Analisis data berbasis *machine learning* berkembang pesat sebagai teknik yang mampu memprediksi penyakit dari pola data pasien. Teknik ini membantu tenaga medis mendeteksi diabetes lebih cepat dan dapat memperkuat sistem pendukung keputusan (*decision support system*) untuk keputusan klinis yang lebih baik (Nurjaya & Rilvani, 2025). Dua algoritma klasifikasi yang banyak dipakai untuk tujuan tersebut adalah Decision Tree dan Naive Bayes. Decision Tree menghasilkan aturan keputusan berbentuk pohon yang mudah ditafsirkan, sedangkan Naive Bayes bekerja cepat dengan asumsi independensi antarfitur (Sihananto & Maulana, 2021).

Sejumlah penelitian telah membandingkan kedua algoritma tersebut pada data diabetes dengan hasil yang beragam. (Maulana et al., 2023) melaporkan akurasi Naive Bayes sebesar 91,56% dan Decision Tree sebesar 87,01%. (Nuryamin & Risyda, 2025) memperoleh akurasi 73,48% untuk Naive Bayes dan 70,87% untuk Decision Tree C4.5 dengan pembagian data 70:30. Sebaliknya, (Ardiansyah et al., 2021) menemukan bahwa C4.5 mencapai akurasi 99,03% dan mengungguli Naive Bayes pada 520 data diabetes dari Kaggle, dan (Permana & Dewi, 2021) juga menyimpulkan bahwa Decision Tree lebih baik. Perbedaan simpulan ini menunjukkan bahwa keunggulan salah satu algoritma bergantung pada data, praproses, dan skema pengujian.

Penelitian terdahulu umumnya menilai kinerja dengan akurasi, presisi, dan *recall*, serta menggunakan pembagian data tunggal atau tidak melaporkan AUC. Padahal, pada data yang tidak seimbang seperti data diabetes, akurasi saja dapat menyesatkan, dan AUC diperlukan untuk menilai kemampuan model memisahkan kelas. Penelitian ini mengisi celah tersebut dengan menggunakan kerangka CRISP-DM, *cross validation*, dan lima metrik evaluasi sekaligus, yaitu *accuracy*, *precision*, *recall*, *F1-score*, dan *Area Under the Curve* (AUC), pada Pima Indians Diabetes Database.

Tujuan penelitian ini adalah mengukur dan membandingkan kinerja Decision Tree dan Naive Bayes dalam memprediksi risiko diabetes menggunakan RapidMiner, kemudian menentukan algoritma yang paling sesuai untuk dataset tersebut. Hasilnya diharapkan menjadi rujukan pemilihan algoritma dalam perancangan DSS skrining awal diabetes.

2. Metode Penelitian

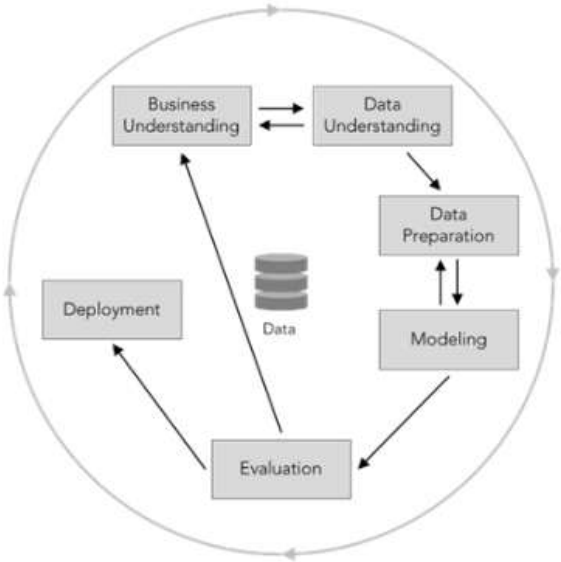
Penelitian ini menerapkan metode kuantitatif berbasis data mining untuk mengevaluasi serta mengomparasikan performa yang ditunjukkan oleh algoritma Decision Tree dan Naive Bayes pada dataset Pima Indians Diabetes. Metodologi dirancang secara sistematis menggunakan RapidMiner sebagai alat utama analisis.

2.1. Tahapan Penelitian

Penelitian ini menggunakan kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data Mining*), yaitu standar penambangan data yang diinisiasi sejak 1996 untuk merumuskan strategi penyelesaian masalah, baik pada riset maupun industri (Fitriani et al., 2022). Kerangka ini terdiri atas enam tahap, yaitu *business understanding, data understanding, data preparation, modeling, evaluation,* dan *deployment* (Mubarak et al., 2024).

Pada tahap *business understanding*, tujuan penelitian ditetapkan, yaitu membandingkan kemampuan pemodelan Decision Tree dan Naive Bayes dalam mengestimasi risiko diabetes. Tahap *data understanding* dan *data preparation* mencakup pengenalan dataset, pemeriksaan kualitas data, dan transformasi data. Tahap *modeling* membangun kedua model di RapidMiner, tahap *evaluation* mengukur kinerja model, dan tahap *deployment* berupa rekomendasi algoritma terbaik sebagai bahan perancangan DSS. Tahapan tersebut diperlihatkan pada Gambar 1.

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 1. Model CRISP-DM (Haiban & Salahuddin, 2025)

2.2. Dataset

Data yang digunakan adalah Pima Indians Diabetes Database dari Kaggle, yang semula dikumpulkan oleh *National Institute of Diabetes and Digestive and Kidney Diseases* (NIDDK). Dataset berisi 768 catatan pasien perempuan keturunan Pima Indian berusia minimal 21 tahun yang bermukim di sekitar Phoenix, Arizona, Amerika Serikat. Setiap catatan memiliki delapan atribut prediktor, yaitu jumlah kehamilan (*Pregnancies*), kadar glukosa plasma 2 jam setelah uji toleransi glukosa oral (*Glucose*), tekanan darah 10008iastolic (*BloodPressure*), ketebalan lipatan kulit trisep (*SkinThickness*), kadar insulin serum 2 jam (*Insulin*), indeks massa tubuh (*BMI*), fungsi silsilah diabetes (*DiabetesPedigreeFunction*), dan usia (*Age*).

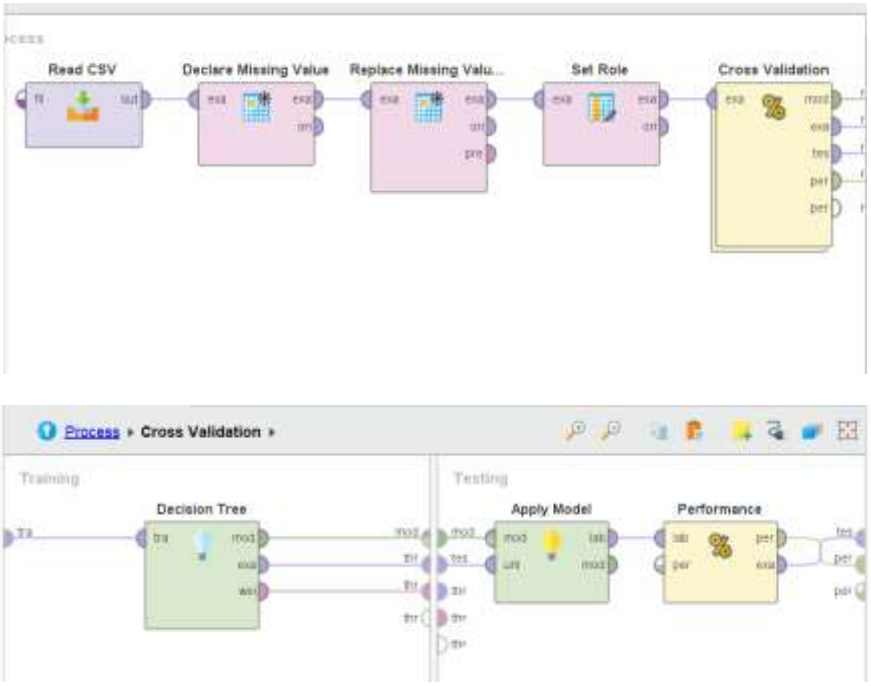
Atribut *Outcome* menjadi label kelas dengan nilai 1 untuk pasien terdiagnosis diabetes sebanyak 268 data (34,9%) dan nilai 0 untuk pasien tidak terdiagnosis diabetes sebanyak 500 data (65,1%). Dengan demikian, kelas diabetes diperlakukan sebagai kelas positif dalam seluruh perhitungan metrik pada penelitian ini.

2.3. Persiapan Data

Pemeriksaan kualitas data menunjukkan adanya nilai nol pada atribut yang secara biologis tidak mungkin bernilai nol, yaitu *glucose*, *BloodPressure*, *SkinThickness*, *insulin*, dan BMI. Nilai nol tersebut ditafsirkan sebagai data hilang. Penanganannya dilakukan dengan dua operator berurutan di RapidMiner. Operator *Declare Missing Values* mengubah nilai nol menjadi nilai hilang, kemudian operator *Replace Missing Values* mengisinya dengan rata-rata (*mean)* atribut yang bersangkutan. Imputasi *mean* dipilih karena distribusi sebagian besar atribut cenderung mendekati normal. Selanjutnya, atribut *Outcome* dikonversi dari numerik menjadi binominal ("*negative*" dan "*positive*") dan ditetapkan sebagai label melalui operator *Set Role*.

Alur proses di RapidMiner terdiri atas operator *Read CSV*, *Declare Missing Values*, *Replace Missing Values*, *Set Role*, dan *Cross Validation*, seperti terlihat pada Gambar 2. Di dalam *Cross Validation* terdapat sisi pelatihan yang berisi operator *Decision Tree* atau *Naive Bayes*, dan sisi pengujian yang berisi operator *Apply Model* dan *Performance*. Alur yang sama dipakai untuk kedua algoritma sehingga perbandingan berlangsung adil.

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 2. Alur proses pada RapidMiner

2.4. Algoritma Decision Tree

Decision Tree membagi data secara rekursif berdasarkan atribut yang paling informatif. Ukuran ketidakmurnian data dihitung dengan entropi pada persamaan (1), dengan *S* adalah himpunan data yang dianalisis, *n* jumlah kelas, dan p_i proporsi data kelas ke-*i* terhadap *S* (Oktavianti et al., 2024).

$$Entropy\ (S)\ =\ \sum_{i=1}^{n} - pi * log2pi \tag{1}$$

Penurunan entropi akibat pemisahan oleh atribut *A* dinyatakan sebagai *gain* pada persamaan (2), dengan S_i adalah subset hasil partisi ke-*i*, $|S_i|$ jumlah data pada subset tersebut, dan $|S|$ jumlah seluruh data. Karena *gain* cenderung memihak atribut dengan banyak partisi, digunakan pula informasi pemisahan (*SplitInfo*) pada persamaan (3) dan rasio *gain* pada persamaan (4)

$$Gain\ (S, A) = Entropy\ (S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy\ (Si) \tag{2}$$

$$\text{SplitInfo}(S, A) = -\sum_{i=1}^{n} \frac{|S_i|}{|S|} \cdot \log_2\left(\frac{|S_i|}{|S|}\right) \tag{3}$$

$$\text{GainRatio}(S, A) = \frac{\text{Gain}(S,A)}{\text{SplitInfo}(S,A)} \tag{4}$$

Pada RapidMiner, kriteria pemisahan diatur menggunakan *accuracy*, kedalaman maksimum pohon 10, ukuran daun minimum 2, dan *pruning* diaktifkan untuk mencegah *overfitting*. Selain itu, proses pembentukan pemisahan ditelusuri secara komputasional dengan *Microsoft Excel*. Ambang (*threshold*) tiap atribut numerik ditetapkan sebagai rata-rata dua nilai berurutan setelah data diurutkan, kemudian dihitung entropi, *gain*, *split info*, rasio *gain*, dan akurasi pemisahan.

2.5. Algoritma Naïve Bayes

Naive Bayes mengklasifikasikan data berdasarkan Teorema Bayes pada persamaan (5), dengan *P(c|x)* probabilitas kelas *c* bila diberikan data *x*, *P(x|c)* kemungkinan (likelihood) data *x* pada kelas *c*, *P(c)* probabilitas awal (*prior*) kelas *c*, dan *P(x)* probabilitas kemunculan data *x* secara umum (Sucahyo & Riskiono, 2026).

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

$$P(c|x) = \frac{P(x|c)\cdot P(c)}{P(x)} \quad (5)$$

Karena seluruh atribut bersifat numerik kontinu, digunakan varian *Gaussian* dengan *likelihood* pada persamaan (6), dengan x_i nilai atribut ke-*i*, μ_c rata-rata atribut pada kelas *c*, dan σ_c simpangan baku atribut pada kelas *c*.

$$P(x_i|c) = \frac{1}{\sqrt{2\pi\sigma_c{}^2}} \cdot exp\left(-\frac{(x_i-\mu_c)^2}{2\sigma_c{}^2}\right) \quad (6)$$

2.6. Evaluasi Kinerja Model

Kinerja kedua model diuji dengan 10-*fold validation*, di mana data dibagi acak menjadi 10 subset dengan ukuran hampir sama, satu subset menjadi data uji dan sembilan lainnya menjadi data latih, diulang 10 kali sehingga setiap subset menjadi data uji tepat satu kali. Hasil pengujian diringkas dalam *confusion matrix* yang memuat *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN) (Romadloni et al., 2022).Metrik dihitung dengan persamaan (7) sampai (10) (Hakim et al., 2025).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

$$F1-Score = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (10)$$

Kinerja klasifikasi diukur dengan empat metrik turunan *confusion matrix*. Akurasi menyatakan proporsi prediksi yang benar terhadap seluruh data uji. *Precision* menyatakan ketepatan prediksi positif, yaitu proporsi prediksi positif yang benar terhadap seluruh prediksi positif. *Recall* menyatakan kemampuan model mengenali kasus positif aktual, sedangkan *F1-score* merupakan rataan harmonik *precision* dan *recall* yang menyeimbangkan keduanya.

Selain keempat metrik tersebut, kemampuan model membedakan pasien diabetes dan non-diabetes dinilai dengan *Area Under the Curve* (AUC). Kurva ROC menggambarkan hubungan antara *true positive rate* dan *false positive rate* pada berbagai ambang klasifikasi, dan AUC adalah luas area di bawah kurva tersebut. Semakin mendekati 1, semakin baik kemampuan model memisahkan kedua kelas (Wilujeng et al., 2023). Nilai AUC diinterpretasikan menurut lima kategori pada Tabel 1 (Fakih et al., 2025).

Tabel 1. Kategori nilai AUC

Rentang AUC	Kategori
0,90–1,00	Sangat baik
0,80–0,90	Baik
0,70–0,80	Cukup baik
0,60–0,70	Kurang baik
0,50–0,60	Gagal

3. Hasil dan Pembahasan

3.1. Kondisi Data dan Hasil Praproses

Pemeriksaan data menemukan bahwa atribut *Insulin* memiliki nilai nol terbanyak, yaitu 374 data (48,70%), diikuti *SkinThickness* sebanyak 227 data (29,56%), seperti terlihat pada Tabel 2. Atribut *BloodPressure*, *BMI*, dan *Glucose* hanya memiliki nilai nol dalam jumlah kecil. Meskipun proporsi data hilang pada *Insulin* dan *SkinThickness* besar, kedelapan atribut tetap dipertahankan karena memiliki relevansi klinis. Nilai nol pada atribut *Pregnancies* (111 data) tidak diperlakukan sebagai data hilang karena nol kehamilan merupakan nilai yang sah.

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tabel 2. Hasil Identifikasi dan Imputasi Nilai Nol pada Dataset Diabetes

Atribut	Jumlah nol	Presentase	Nilai Pengganti (mean)
Glucose	5	0,65%	122
BloodPressure	35	4,56%	72
SkinThickness	227	29,56%	29
Insulin	374	48,70%	156
BMI	11	1,43%	32,457

Setelah imputasi, seluruh data lengkap dan siap dipakai. Pendekatan ini sederhana dan menjaga jumlah data tetap 768, tetapi dapat mengecilkan variasi pada atribut yang proporsi imputasinya tinggi, sehingga perlu diperhatikan dalam menafsirkan hasil.

3.2. Hasil Pembentukan Decision Tree

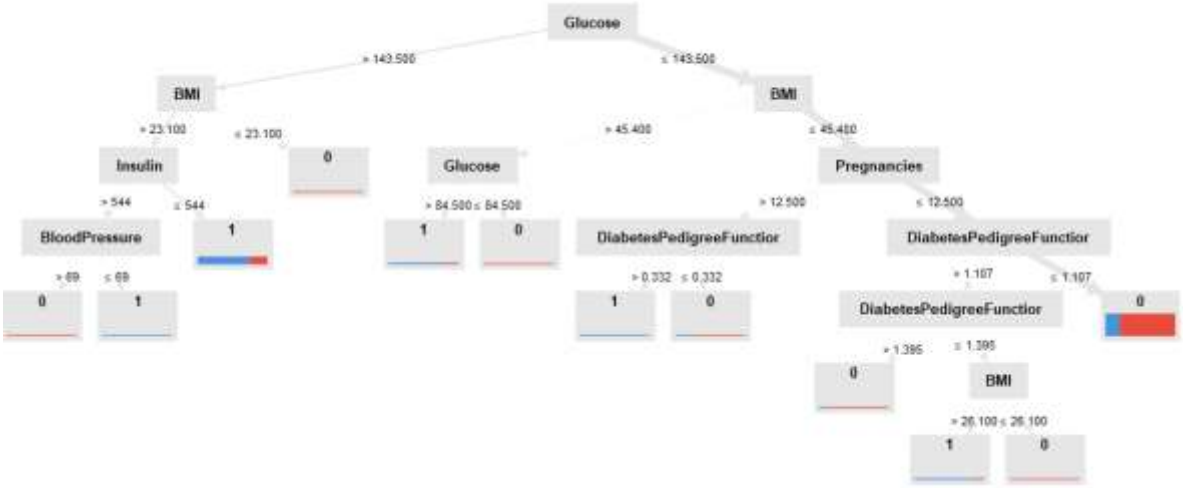
Hasil penelusuran komputasional pada Tabel 3 menunjukkan bahwa atribut *Glucose* dengan ambang 143,5 menghasilkan *gain* tertinggi (0,1232), rasio *gain* tertinggi (0,1586), dan akurasi pemisahan tertinggi (75,00%) dibandingkan tujuh atribut lainnya. Atribut ini membagi data menjadi 592 data dengan *Glucose* ≤ 143,5 (142 diabetes, 450 tidak diabetes) dan 176 data dengan *Glucose* > 143,5 (126 diabetes, 50 tidak diabetes). Oleh karena itu, *Glucose* ditetapkan sebagai simpul akar pohon keputusan.

Tabel 3. Hasil Perhitungan Pemisahan Terbaik Tiap Atribut

Atribut	Threshold	Kondisi	Jumlah (S)	Kelas 1 (Diabet)	Kelas 0 (Tidak Diabet)	Entropy	Gain	Split Info	Gain Ratio	Accuracy
Total (S) seluruh data			768	268	500	0,9331				
Pregnancies	6,5	≤ 6,5	599	173	426	0,8672	0,0392	0,7603	0,0515	67,84%
	6,5	> 6,5	169	95	74	0,9888				
Glucose	143,5	≤ 143,5	592	142	450	**0,7948**	**0,1232**	**0,7766**	**0,1586**	**75,00%**
	143,5	> 143,5	176	126	50	0,8610				
Blood Pressure	97	≤ 97	749	257	492	0,9278	0,0040	0,1673	0,0240	65,49%
	97	> 97	19	11	8	0,9819				
Skin Thickness	43,5	≤ 43,5	726	246	4 80	0,9237	0,0053	0,3060	0,0175	65,36%
	43,5	> 43,5	42	22	20	0,9984				
Insulin	204,5	≤ 204,5	684	222	462	0,9093	0,0147	0,4980	0,0294	66,15%
	204,5	> 204,5	84	46	38	0,9934				
BMI	40,85	≤ 40,85	683	215	468	0,8986	0,0282	0,5020	0,0562	67,84%
	40,85	> 40,85	85	53	32	0,9555				
Diabetes PedigreeF	1,1075	≤ 1,1075	727	243	484	0,9192	0,0115	0,3006	0,0382	66,28%
	1,1075	> 1,1075	41	25	16	0,9650				
Age	40,5	≤ 40,5	574	166	408	0,8677	0,0325	0,8154	0,0399	66,41%
	40,5	> 40,5	194	102	92	0,9981				

Dominasi *Glucose* konsisten dengan sebaran data per kelas, karena rata-rata *Glucose* pasien diabetes (142,17 mg/dL) jauh di atas pasien non-diabetes (110,71 mg/dL). Pohon keputusan yang dihasilkan RapidMiner pada Gambar 3 memperlihatkan pola yang sama dan dapat dibaca langsung sebagai aturan jika-maka, sehingga mudah dijelaskan kepada pengguna non-ahli.

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 3. Pohon Keputusan pada RapidMiner

3.3. Hasil Naive Bayes

Tahap pertama adalah menghitung probabilitas *prior* untuk mengetahui proporsi masing-masing kelas dalam dataset. Dataset yang digunakan terdiri dari 768 data, dengan 500 data Outcome 0 dan 268 data Outcome 1. Perhitungan probabilitas *prior* dilakukan sebagai berikut:

$$\text{P (Outcome 0)} = \frac{Jumlah\ Outcome = 0}{Total\ Data} = \frac{500}{768} = 0{,}65104$$

$$\text{P (Outcome 1)} = \frac{Jumlah\ Outcome = 1}{Total\ Data} = \frac{268}{768} = 0{,}34896$$

Tahap berikutnya menghitung rata-rata (*mean*) dan simpangan baku setiap atribut pada masing-masing kelas dengan *Microsoft Excel*, memakai fungsi AVERAGEIFS untuk rata-rata dan STDEV.P untuk simpangan baku. Hasilnya disajikan pada Tabel 4.

Tabel 4. Hasil Perhitungan Mean Dan Standar Deviasi

Atribut	Mean | C=0	Stdev | C=0	Mean | C=1	Stdev | C=1
Pregnancies	3,298	3,014	4,866	3,734
Glucose	110,710	24,692	142,166	29,487
BloodPressure	70,935	11,919	75,147	11,923
SkinThickness	27,769	8,551	31,737	8,631
Insulin	142,211	75,388	180,432	95,533
BMI	30,888	6,498	35,385	10,948
DiabetesPedigreeFunction	0,430	0,299	0,551	0,372
Age	31,190	11,656	37,067	10,948

Rata-rata *Glucose* pasien diabetes (142,17 mg/dL) jauh di atas pasien non-diabetes (110,71 mg/dL) dengan sebaran yang relatif terpisah, sehingga *Glucose* memiliki daya pembeda kelas paling kuat. Temuan ini konsisten dengan terpilihnya *Glucose* sebagai simpul akar pada *Decision Tree*. Nilai rata-rata dan simpangan baku tersebut kemudian dipakai untuk menaksir *likelihood* Gaussian setiap atribut dengan fungsi NORM.DIST pada mode densitas probabilitas.

Penghitungan diilustrasikan pada data uji pertama, yaitu 6, 148, 72, 35, 156, 33,6, 0,627, dan 50 untuk kedelapan atribut. Tabel 5 memuat *likelihood* tiap atribut pada kedua kelas. Karena merupakan nilai densitas, *likelihood* dapat melebihi 1, seperti pada *DiabetesPedigreeFunction*. Perkalian kedelapan *likelihood* menghasilkan *P(X|C)* sebesar $1{,}456423 \times 10^{-12}$ untuk kelas 0 dan $5{,}213246 \times 10^{-12}$ untuk kelas 1. Setelah dikalikan *prior*, skor posterior kelas 0 sebesar $9{,}481920 \times 10^{-13}$ lebih kecil daripada kelas 1 sebesar $1{,}819206 \times 10^{-12}$, sehingga data tersebut diklasifikasikan sebagai diabetes, sesuai dengan label aktualnya.

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tabel 5. Hasil Perhitungan *Likelihood Gaussian* Dan Skor Posterior Pada Data Uji Pertama

Atribut	Nilai uji	*Likelihood* C=0	*Likelihood* C=1
Pregnancies	6	0,0886	0,1020
Glucose	148	0,0052	0,0133
BloodPressure	72	0,0333	0,0323
SkinThickness	35	0,0326	0,0430
Insulin	156	0,0052	0,0040
BMI	33,6	0,0563	0,0360
DiabetesPedigree	0,627	1,0737	1,0509
Age	50	0,0093	0,0181
P(X|C)		$1{,}456423 \times 10^{-12}$	$5{,}213246 \times 10^{-12}$
Prior		0,65104	0,34896
Posterior		$\mathbf{9{,}481920 \times 10^{-13}}$	$\mathbf{1{,}819206 \times 10^{-12}}$

3.4 Evaluasi dan Perbandingan Kinerja

Tabel 6 memperlihatkan *confusion matrix* gabungan seluruh lipatan, dengan kelas diabetes sebagai kelas positif. *Decision Tree* memperoleh 122 TP, 443 TN, 57 FP, dan 146 FN, sedangkan *Naive Bayes* memperoleh 160 TP, 412 TN, 88 FP, dan 108 FN. *Naive Bayes* mengenali 38 pasien diabetes lebih banyak daripada *Decision Tree*, tetapi menghasilkan 31 alarm palsu lebih banyak.

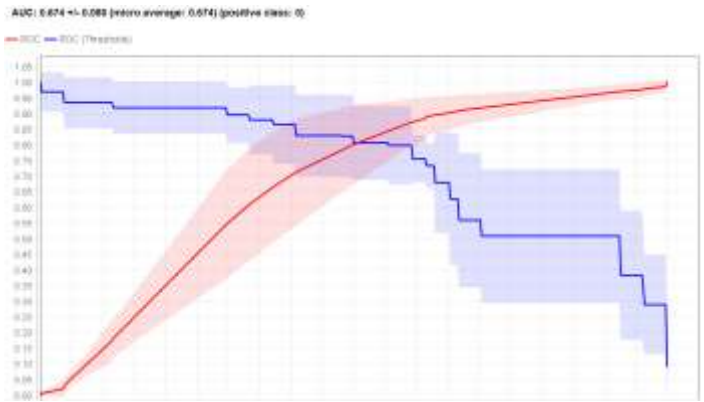
Tabel 6. Confusion Matrix

Algoritma	TP	TN	FP	FN
Decision Tree	122	443	57	146
Naive Bayes	160	412	88	108

Dari nilai tersebut, *Decision Tree* memperoleh akurasi 73,57%, *precision* 68,16%, *recall* 45,52%, dan *F1-score* 54,59%, sedangkan *Naive Bayes* memperoleh akurasi 74,48%, *precision* 64,52%, *recall* 59,70%, dan *F1-score* 62,02%. Nilai AUC *Decision Tree* sebesar 0,674 termasuk kategori kurang baik, sedangkan *Naive Bayes* sebesar 0,817 termasuk kategori baik. Nilai AUC beserta kategorinya dirangkum bersama metrik lain pada Tabel 7.

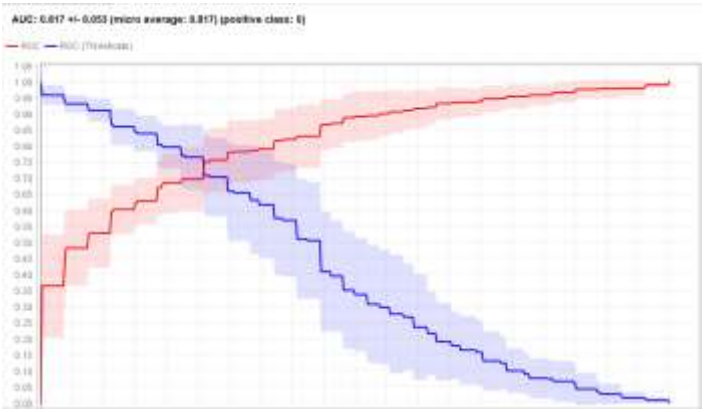
Tabel 7. Perbandingan kinerja Decision Tree dan Naive Bayes

Metrik	***Decision Tree***	***Naive Bayes***	**Algoritma Terbaik**
Accuracy	73,57%	74,48%	*Naive Bayes*
Precision	68,16%	64,52%	*Decision Tree*
Recall	45,52%	59,70%	*Naive Bayes*
F1-Score	54,59%	62,02%	*Naive Bayes*
AUC	0,674	0,817	*Naive Bayes*



Gambar 4. Hasil AUC Decision Tree

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 5. Hasil AUC Naive Bayes

3.5. Pembahasan

Hasil pada Tabel 7 menunjukkan bahwa Naive Bayes lebih baik daripada Decision Tree pada akurasi (74,48%), recall (59,70%), F1-score (62,02%), dan AUC (0,817). Selisih akurasinya kecil, hanya 0,91 poin, sehingga keunggulan Naive Bayes terutama terletak pada kemampuannya mengenali pasien diabetes. Model ini menemukan 38 pasien diabetes lebih banyak daripada Decision Tree (160 berbanding 122), sedangkan nilai AUC 0,817 menunjukkan kemampuan memisahkan kedua kelas yang baik. Pada Decision Tree, AUC sebesar 0,674 hanya termasuk kategori kurang baik.

Keunggulan tersebut disertai kompromi. Naive Bayes menghasilkan 31 false positive lebih banyak (88 berbanding 57), sehingga precision-nya lebih rendah, yaitu 64,52% berbanding 68,16%. Kompromi ini dapat diterima pada skrining awal karena pasien diabetes yang terlewat lebih berisiko daripada pasien sehat yang diminta menjalani pemeriksaan lanjutan. Dengan pertimbangan tersebut, Naive Bayes dipilih sebagai algoritma yang lebih sesuai untuk memprediksi risiko diabetes pada dataset ini.

4. Kesimpulan

Penelitian ini membandingkan kinerja Decision Tree dan Naive Bayes dalam memprediksi risiko diabetes pada Pima Indians Diabetes Database menggunakan RapidMiner dan validasi silang 10 lipatan. Naive Bayes memperoleh akurasi 74,48%, recall 59,70%, F1-score 62,02%, dan AUC 0,817, lebih tinggi daripada Decision Tree yang memperoleh akurasi 73,57%, recall 45,52%, F1-score 54,59%, dan AUC 0,674. Decision Tree hanya unggul pada precision, yaitu 68,16% berbanding 64,52%. Dengan demikian, Naive Bayes merupakan algoritma yang lebih sesuai untuk memprediksi risiko diabetes pada dataset ini, yang ditunjukkan oleh recall dan AUC yang lebih tinggi. Penelitian selanjutnya disarankan menerapkan teknik penyeimbangan data, seperti SMOTE, mengevaluasi algoritma klasifikasi lain, seperti K-Nearest Neighbor, Random Forest, dan Support Vector Machine, serta menggunakan dataset dari populasi yang lebih beragam untuk mengetahui konsistensi dan meningkatkan kemampuan generalisasi model.

Reference

1. Ardiansyah, M., Sunyoto, A., & Luthfi, E. T. (2021). *Analisis Perbandingan Akurasi Algoritma Naive Bayes dan C4.5 untuk Klasifikasi Diabetes*. *5*(2), 147–156. https://doi.org/10.29408/edumatic.v5i2.3424
2. Depari, A. D. S., Kirana, C. C., Oktariana, C. N., Akbar, F., & Fathoni. (2025). *PREDIKSI RISIKO DIABETES DENGAN METODE NAIVE BAYES : IDENTIFIKASI FAKTOR RISIKO UTAMA DAN EVALUASI AKURASI MODEL*. *9*(4), 6372–6377. https://doi.org/10.36040/jati.v9i4.14078
3. Fakih, A., Hamzami, M. A., Hadianto, M. R., & Alifah, N. I. S. (2025). Perbandingan Akurasi Algoritma C4.5 dan K-NN Untuk Prediksi Kelulusan Mahasiswa Penerima Beasiswa. *Jurnal Komputer Antartika*, *3*(1), 18–25. https://doi.org/10.70052/jka.v3i1.623
4. Febrian, M. G., Ferdiansyah, R., Nugraha, E. A., Satriatama, D., & Rajnaparamitha, K. (2024). *PREDIKSI RISIKO DIABETES MENGGUNAKAN ALGORITMA DECISION TREE DENGAN APLIKASI RAPID MINER*. *November*, 14–24. https://doi.org/10.9644/scp.v1i1.332
5. Fitriani, M., Nama, G. F., Informatika, T., & Lampung, U. (2022). *IMPLEMENTASI ASSOCIATION RULE DENGAN ALGORITMA APRIORI PADA DATA PEMINJAMAN BUKU UPT PERPUSTAKAAN UNIVERSITAS LAMPUNG MENGGUNAKAN METODOLOGI CRISP-DM*. *10*(1), 41–49. https://doi.org/10.23960/jitet.v10i1.2263
6. Haiban, M. F. N., & Salahuddin, S. N. (2025). *Ekstraksi Wawasan Dari Platform X (Twitter): Aplikasi Crisp-Dm Dan Visualisasi*

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)

Tableau Dalam Analisis Sentimen Debat Capres Dan Cawapres Pertama 2024 Di Indonesia. *8*, 191–200. https://doi.org/10.31539/intecoms.v8i1.13311

7. Hakim, L., Sobri, A., Sunardi, L., & Nurdiansyah, D. (2025). *Prediksi penyakit jantung berbasis mesin learning dengan menggunakan metode k-nn*. *07*(02), 14–20. https://doi.org/10.32502/digital.v7i2.9429
8. Maulana, R., Narasati, R., Herdiana, R., Hamonangan, R., & Anwar, S. (2023). *KOMPARASI ALGORITMA DECISION TREE DAN NAIVE BAYES DALAM KLASIFIKASI PENYAKIT DIABETES*. *7*(6), 3865–3870. https://doi.org/10.36040/jati.v7i6.8265
9. Mubarak, R., Hanafi, M., & Sasongko, D. (2024). Komparasi Performa Naive Bayes Gaussian dan K-NN Untuk Prediksi Kelulusan Mahasiswa dengan CRISP-DM. *Media Online)*, *4*(6), 2982–2991. https://doi.org/10.30865/klik.v4i6.1924
10. Nurjaya, & Rilvani, E. (2025). *PERBANDINGAN AKURASI NAÏVE BAYES DECISION TREE UNTUK KLASIFIKASI RISIKO DIABETES*. *10*(4). https://doi.org/10.2238/g2afsv55
11. Nuryamin, Y., & Risyda, F. (2025). *Analisis Prediksi Penyakit Diabetes Menggunakan Metode Decision Tree C4 . 5 Dan Naive Bayes*. *5*, 234–242. https://doi.org/10.35968/jsi.v12i2.1549
12. Oktavianti, E., Agustin, M., & Sari, R. (2024). Implementasi Algoritma Decision Tree Dengan Fitur Seleksi Weight By Information Gain. *Multinetics*, *9*(2), 118–126. https://doi.org/10.32722/multinetics.v9i2.5715
13. Permana, B. A. C., & Dewi, I. K. (2021). *Komparasi Metode Klasifikasi Data Mining Decision Tree dan Naive Bayes Untuk Prediksi Penyakit Diabetes*. *4*(1), 63–69. https://doi.org/10.29408/jit.v4i1.2994
14. Romadloni, P. L., Kusuma, B. A., & Baihaqi, W. M. (2022). Komparasi Metode Pembelajaran Mesin Untuk Implementasi Pengambilan Keputusan Dalam Menentukan Promosi Jabatan Karyawan. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *6*(2), 622–628. https://doi.org/10.36040/jati.v6i2.5238
15. Sihananto, A. N., & Maulana, H. (2021). *Studi Literatur Tentang Performa Naïve Bayes dalam Klasifikasi Data*. *2*, 132–135. https://doi.org/10.33005/santika.v2i0.134
16. Sucahyo, B. D., & Riskiono, S. D. (2026). *Perbandingan Algoritma K-Nearest Neighbor dan Naive Bayes Untuk Klasifikasi Status Penjualan Furniture dengan Data Historis*. *7*(2). https://doi.org/10.47065/josh.v7i2.8675
17. Wilujeng, D. T., Fatekurohman, M., & Tirta, I. M. (2023). Analisis Risiko Kredit Perbankan Menggunakan Algoritma K-Nearest Neighbor dan Nearest Weighted K-Nearest Neighbor. *Indonesian Journal of Applied Statistics*, *5*(2), 142–148. https://doi.org/10.13057/ijas.v5i2.58426

DOI: https://doi.org/10.69693/ijmst.v4i3.14364
Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)