

KLASIFIKASI PENYAKIT PREEKSLAMISIA PADA IBU HAMIL MENGGUNAKAN PERBANDINGAN ALGORITMA *MACHINE LEARNING*

Novita Ranti Muntiari ^[1]; Kharis Hudaiby Hanif ^[2]; Syamsiah ^[3]; Rokaya ^[4]

[1,3,4] Politeknik Kaltara Program Studi Promosi Kesehatan

[2] Universitas Borneo Tarakan Program Studi Teknik Komputer

Email: novitarantimuntiari@gmail.com

INFO ARTIKEL	INTISARI
Diajukan : 12 Juli 2025	Preeklampsia merupakan salah satu komplikasi kehamilan yang berpotensi mengancam keselamatan ibu dan janin apabila tidak terdeteksi secara dini. Kondisi ini memerlukan pendekatan analisis yang cepat dan tepat, terutama pada fasilitas kesehatan dengan keterbatasan sumber daya. Penelitian ini memanfaatkan pendekatan <i>machine learning</i> untuk membangun model klasifikasi risiko preeklampsia melalui analisis data kesehatan maternal. Data diperoleh dari dataset publik Kaggle dan diolah melalui beberapa tahapan, meliputi pra-proses data, pembagian data menjadi 80% untuk pelatihan dan 20% untuk pengujian, serta penerapan beberapa algoritma klasifikasi. Model yang dibandingkan dalam penelitian ini meliputi <i>Logistic Regression</i> , <i>K-Nearest Neighbor</i> , <i>Decision Tree</i> , <i>Random Forest</i> , dan <i>Support Vector Machine</i> . Hasil evaluasi menunjukkan bahwa algoritma <i>Decision Tree</i> dan <i>Random Forest</i> memiliki performa paling baik dengan akurasi di atas 87%, sedangkan <i>Logistic Regression</i> menjadi model dengan akurasi terendah. Temuan ini menunjukkan bahwa metode berbasis pohon keputusan lebih sesuai untuk memetakan pola risiko preeklampsia berdasarkan variabel klinis. Penelitian ini diharapkan dapat mendukung pengembangan dalam mendeteksi risiko preeklampsia secara lebih efektif.
Diterima : 20 Agustus 2025	
Diterbitkan: 27 Desember 2025	
Kata Kunci : <i>Preeklampsia</i> , Klasifikasi, <i>Decision Tree</i> , KNN, <i>Machine Learning</i>	

I. PENDAHULUAN

Preeklampsia merupakan salah satu komplikasi kehamilan yang paling sering ditemukan dan menjadi penyebab penting meningkatnya angka kesakitan serta kematian pada ibu dan janin, terutama di negara berkembang. Kondisi ini ditandai oleh peningkatan tekanan darah disertai gangguan fungsi organ, yang biasanya muncul setelah usia kehamilan 20 minggu (Fitriana et al., 2021). Apabila tidak terdeteksi sejak dini, preeklampsia dapat berkembang menjadi eklampsia, gangguan ginjal, gangguan hati, bahkan kematian (Hanif et al., 2022). Kompleksitas gejala serta progresi penyakit yang sering kali terjadi secara tiba-tiba membuat preeklampsia membutuhkan proses identifikasi yang cepat, akurat, dan berkelanjutan. Di sisi lain, keterbatasan tenaga kesehatan, variasi kualitas pemeriksaan, dan heterogenitas kondisi ibu hamil menyebabkan deteksi dini secara manual sering tidak optimal (Rawansyah et al., 2020).

Kemajuan teknologi data science, khususnya bidang *machine learning*, membuka peluang baru dalam mendukung proses penilaian risiko preeklampsia. Algoritma *machine learning* mampu mempelajari pola dan hubungan antar variabel klinis berdasarkan data yang tersedia, sehingga dapat menghasilkan prediksi yang lebih objektif

dan konsisten. Pendekatan ini sangat bermanfaat untuk membantu tenaga kesehatan, terutama dalam proses skrining awal, pemantauan ibu hamil berisiko, serta pengambilan keputusan lebih cepat di layanan kesehatan primer maupun rumah sakit. *Machine learning* semakin banyak dimanfaatkan karena kemampuannya dalam mengolah data klinis yang beragam dan jumlahnya besar (Aisha et al., 2024), sehingga dapat membantu mempercepat analisis dan meningkatkan ketepatan diagnosis (Tjahjadi & Ramli, 2020).

Algoritma *machine learning* salah satunya yaitu *naive bayes*, *random forest*, *Support Vector Machine (SVM)*, *logistic regression* (Muntiari et al., 2021).

Penelitian (Muntiari & Hanif, 2022) menggunakan *machine learning* dengan bertujuan mengklasifikasikan jenis kanker payudara. Tujuh algoritma yang digunakan meliputi *decision tree*, *KNN*, *logistic regression*, dan *random forest*, dan *SVM* dan lainnya. Hasilnya 95% nilai akurasi tertinggi dari setiap algoritma *machine learning*. Penelitian yang dilakukan oleh (Chazar & Erawan, 2020) bertujuan mendiagnosis kanker payudara menggunakan *SVM*. Hasil penelitian mampu menentukan sel hidup kanker bersifat ganas atau jinak.

Perbandingan penelitian ini menggunakan perbandingan algoritma dari *machine learning*

yaitu SVM, KNN, *decision tree*, *random forest* dan *logistic regresion*. Penelitian ini bertujuan menemukan model klasifikasi paling optimal yang mampu memberikan hasil prediksi risiko preeklampsia secara lebih tepat dan dapat diandalkan.

Berdasarkan permasalahan penelitian di atas, maka penelitian ini membuat sebuah model klasifikasi risiko preeklampsia pada ibu hamil dengan memanfaatkan beberapa algoritma *machine learning* guna mengetahui metode mana yang memberikan performa paling optimal dalam mendeteksi tingkat risiko. Penelitian ini tidak hanya membangun satu model, tetapi juga membandingkan kinerja beberapa algoritma, seperti *Logistic Regression*, *Decision Tree*, *Random Forest*, *K-Nearest Neighbor*, dan *Support Vector Machine*, sehingga dapat diperoleh algoritma terbaik yang mampu memberikan prediksi paling akurat terhadap kondisi ibu hamil.

II. BAHAN DAN METODE



Gambar 1. Alur Penelitian

Berdasarkan Gambar 1 alur penelitian ini terdiri dari empat tahap utama yang saling berurutan dan membentuk proses sistematis dalam membangun model klasifikasi risiko preeklampsia berbasis *machine learning*.

Tahap pertama adalah pengumpulan Data, yaitu proses memperoleh dataset maternal *health* dari sumber terbuka seperti Kaggle. Pada tahap ini, seluruh fitur klinis seperti usia, tekanan darah, kadar gula darah, suhu tubuh, serta denyut nadi dikumpulkan sebagai variabel prediktor.

Setelah data terkumpul, penelitian dilanjutkan pada tahap Pra-proses Data. Tahap ini mencakup proses pembersihan data, normalisasi, transformasi label, dan pembagian data menjadi *training* serta *testing*. Tujuan utama pra-proses adalah memastikan bahwa dataset berada dalam kondisi yang ideal untuk dianalisis oleh model *machine learning* sehingga hasil prediksi lebih akurat dan stabil.

Tahap selanjutnya adalah pemodelan dan klasifikasi, yaitu proses pelatihan berbagai algoritma *machine learning* seperti *Logistic Regression* serta 4 algoritma yang lain. Pada tahap ini, setiap model dilatih menggunakan data *training* untuk mempelajari pola hubungan antar variabel yang dapat digunakan untuk mengklasifikasikan risiko preeklampsia menjadi dua kategori, yaitu *low risk* dan *high risk*.

Tahap terakhir dalam alur penelitian adalah evaluasi model. Pada tahap ini, setiap model diuji menggunakan data *testing* dan dievaluasi

menggunakan metrik seperti akurasi, *precision*, *recall*, *F1-score*, serta *confusion matrix*. Hasil evaluasi digunakan untuk menentukan model terbaik yang memberikan performa paling akurat dan dapat diandalkan dalam mendeteksi risiko preeklampsia pada ibu hamil.

Klasifikasi

Dalam *machine learning*, proses klasifikasi terdiri dari dua tahap utama, yaitu pelatihan model (*training*) dan pengujian model (*testing*). Pada tahap pelatihan, model mempelajari pola dari data latih sehingga mampu mengenali hubungan antara variabel input dan kelas target (Muntiari et al., 2023). Setelah itu, model diuji menggunakan data uji yang tidak pernah dilihat sebelumnya untuk mengevaluasi kemampuan generalisasi model (Rahayu et al., 2021).

Dalam penelitian ini, data dibagi menggunakan teknik *train-test split*, yaitu 80% data digunakan untuk melatih model, sedangkan 20% sisanya digunakan untuk menguji model (Oktavianto & Handri, 2020). Pembagian ini penting untuk memastikan bahwa performa model dapat dinilai secara objektif dan tidak bias terhadap data pelatihan. Pembagian data dilihat pada Gambar 2.



Gambar 2. Pembagian *Machine Learning*

adalah bidang ilmu komputasi yang berfokus pada pengembangan sistem yang mampu mengenali pola dan membuat keputusan berdasarkan data. Konsep utamanya adalah memberi komputer kemampuan untuk belajar melalui pengalaman, tanpa harus diberikan aturan yang dituliskan secara eksplisit oleh manusia. Proses pembelajaran ini memungkinkan sebuah model memahami hubungan antar variabel, menyesuaikan diri dengan data baru, serta memberikan prediksi atau klasifikasi sesuai tujuan yang ditetapkan. Dengan kemampuan tersebut, *machine learning* banyak digunakan pada berbagai bidang seperti kesehatan, finansial, pendidikan, hingga pengolahan citra dan bahasa (Priya & Abbas Rizvi, 2020).

Algoritma *Decision Tree*

adalah metode klasifikasi yang membangun model dalam bentuk struktur pohon yang terdiri atas simpul keputusan dan cabang-cabang aturan. Setiap simpul pada pohon merepresentasikan pengujian terhadap suatu atribut, sedangkan setiap cabang menunjukkan hasil dari pengujian tersebut sampai akhirnya mencapai daun yang berisi kelas prediksi. Proses pembentukan pohon dilakukan dengan memilih atribut yang mampu memisahkan data secara paling informatif sehingga pola

hubungan antar variabel dapat terbaca dengan jelas. Keunggulan utama *Decision Tree* adalah kemampuannya menampilkan proses klasifikasi secara transparan dan mudah dipahami, karena alur penentuan keputusan dapat ditelusuri langkah demi langkah. Selain itu, model ini dapat digunakan pada data numerik maupun kategorik dan tidak memerlukan normalisasi data, sehingga sering menjadi pilihan dalam analisis awal untuk memahami struktur dan pola dalam dataset (Yudha Prawira et al., n.d.).

Algoritma SVM

adalah metode klasifikasi yang bekerja dengan mencari garis pemisah terbaik untuk membedakan dua kelompok data. Prinsip utamanya adalah memilih *hyperplane* yang memiliki jarak terbesar terhadap titik-titik data dari masing-masing kelas. Semakin besar jarak atau margin tersebut, semakin baik kemampuan model dalam membedakan kelas dan mengurangi kesalahan prediksi (Wiratama & Aziz, 2024). SVM juga dapat menangani pola data yang tidak linear dengan memanfaatkan fungsi kernel, sehingga data yang kompleks dapat dipetakan ke ruang berdimensi lebih tinggi agar lebih mudah dipisahkan. Dengan sifatnya yang fokus pada titik-titik batas atau *support vectors*, SVM sering memberikan performa yang kuat meskipun pada dataset dengan ukuran relatif kecil (Wiratama & Aziz, 2024).

Algoritma KNN

adalah metode klasifikasi yang bekerja dengan prinsip kedekatan data. Ketika sebuah data baru akan diprediksi, algoritma ini mencari sejumlah tetangga terdekat sebanyak nilai K yang telah ditentukan berdasarkan jarak tertentu, biasanya jarak *Euclidean*. Keputusan kelas data baru ditentukan dari mayoritas kelas tetangga terdekat tersebut (Ramadhani et al., 2024). KNN tidak membangun model atau persamaan matematis, melainkan menyimpan data pelatihan dan melakukan pencarian tetangga ketika proses prediksi berlangsung. Karena mekanismenya yang sederhana namun efektif, KNN sering digunakan pada berbagai permasalahan klasifikasi yang membutuhkan pendekatan berbasis kemiripan pola (Nurmalasari et al., 2021).

Algoritma *Random Forest*

merupakan metode pembelajaran mesin yang bekerja dengan membangun banyak pohon keputusan secara bersamaan. Setiap pohon dilatih menggunakan sampel data yang dipilih secara acak serta subset fitur tertentu. Ketika proses prediksi dilakukan, seluruh pohon memberikan hasil klasifikasi, lalu keputusan akhir ditentukan melalui proses pemungutan suara mayoritas (Taye, 2023).

Pendekatan ini membuat *Random Forest* lebih stabil, tahan terhadap *overfitting*, dan mampu menangani pola data yang kompleks. Kombinasi banyak pohon keputusan tersebut menjadikan algoritma ini efektif untuk berbagai tugas klasifikasi, termasuk dalam memprediksi risiko kesehatan seperti preeklampsia (Nour & Polat, 2020).

Algoritma *Logistic Regression*

metode statistik yang digunakan untuk memprediksi suatu kejadian dengan keluaran berupa dua kemungkinan, seperti “ya” atau “tidak”, “positif” atau “negatif”, serta kategori biner lainnya. Pendekatan ini bekerja dengan menghubungkan variabel input terhadap peluang terjadinya suatu kelas melalui fungsi logistik atau *sigmoid*. Nilai keluaran yang dihasilkan berada pada rentang 0 hingga 1, sehingga dapat ditafsirkan sebagai probabilitas (Kazerouni et al., 2020). *Logistic Regression* banyak digunakan dalam klasifikasi karena modelnya sederhana, mudah dijelaskan, dan mampu memberikan gambaran seberapa besar pengaruh masing-masing variabel terhadap keputusan klasifikasi (Kazerouni et al., 2020).

Confusion Matrix

sebuah tabel evaluasi yang menunjukkan jumlah prediksi benar dan salah pada setiap kelas dalam sebuah model klasifikasi. Matriks ini menampilkan hubungan antara label sebenarnya dan hasil prediksi model, sehingga kesalahan klasifikasi dapat terlihat dengan jelas (Makmur Mulyono1, Elvia Budianita2, Alwis Nazir3, 2023). Melalui susunan nilai *true positive*, *true negative*, *false positive*, dan *false negative*, *confusion matrix* membantu peneliti memahami pola kesalahan serta menilai kemampuan model dalam membedakan setiap kategori kelas. Dengan demikian, *confusion matrix* berperan sebagai alat analisis penting untuk mengukur ketepatan dan keandalan model klasifikasi (Fauzan Adzim et al., 2023).

Evaluasi

Tabel 1. Tabel *Confusion Matrix*

	Prediction Class	
	+	-
+	TP	FN
-	FP	TN

Tabel 1 menyajikan *confusion matrix* yang digunakan untuk menilai performa setiap algoritma klasifikasi. Melalui matriks ini, beberapa metrik evaluasi seperti F1-score, recall, akurasi, dan presisi dapat dihitung berdasarkan rumus masing-masing. Akurasi menjadi salah satu indikator yang paling sering digunakan karena menunjukkan proporsi prediksi yang benar dibandingkan dengan keseluruhan data yang diuji.

Dengan kata lain, akurasi mengukur seberapa besar kemampuan model dalam menghasilkan prediksi yang tepat dari total sampel yang tersedia. Dengan rumus yang digunakan sebagai berikut :

$$F - measure = \frac{2 \times precision \times recall}{precision + recall} \dots\dots\dots(1)$$

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots(2)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots\dots(3)$$

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

III. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan *Maternal Health Risk Data Set* dari Kaggle untuk mengklasifikasikan risiko preeklampsia pada ibu hamil berdasarkan enam parameter klinis: usia, tekanan darah sistolik, tekanan darah diastolik, kadar gula darah, suhu tubuh, dan denyut nadi. Variabel target (*RiskLevel*) awalnya terdiri dari tiga kategori (*Low, Medium, High*), namun dalam penelitian ini disederhanakan menjadi klasifikasi biner, yakni:

1. *Low Risk* = 0
2. *Medium Risk + High Risk* = 1 (*High Risk*)

Total terdapat 1014 data yang kemudian dibagi menggunakan skema *train-test split* sebesar 80% data training (811 data) dan 20% data testing (203 data). Pembagian ini memastikan bahwa model diuji dengan data yang belum pernah dilihat sebelumnya sehingga evaluasi akurasi dapat dilakukan secara objektif.

Untuk meningkatkan performa model berbasis perhitungan jarak seperti *Logistic Regression, KNN*, dan *Support Vector Machine (SVM)*, dilakukan proses *standardisasi* menggunakan *StandardScaler* agar seluruh fitur memiliki skala yang setara. Sementara itu, model berbasis pohon seperti *Decision Tree* dan *Random Forest* tidak memerlukan *standardisasi*.

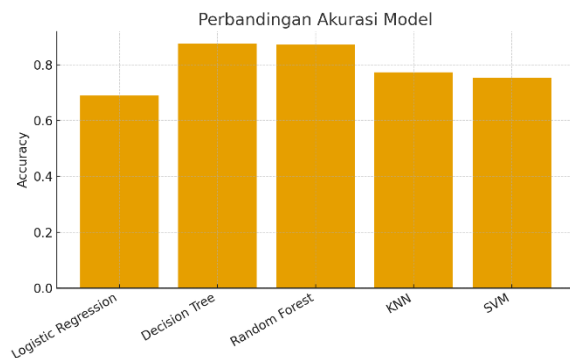
Hasil Evaluasi Model

Evaluasi dilakukan menggunakan *accuracy, precision, recall, F1-score*, serta *confusion matrix* untuk melihat kinerja masing-masing model dalam memprediksi risiko preeklampsia. Berikut ringkasan nilai akurasi masing-masing model seperti pada Tabel 2.

Tabel 2. Nilai Masing-Masing Model

Model	Akurasi
Logistic Regression	68,97 %
Decision Tree	87,68%
Random Forest	87,19 %
KNN	77,34 %
SVM	75,37 %

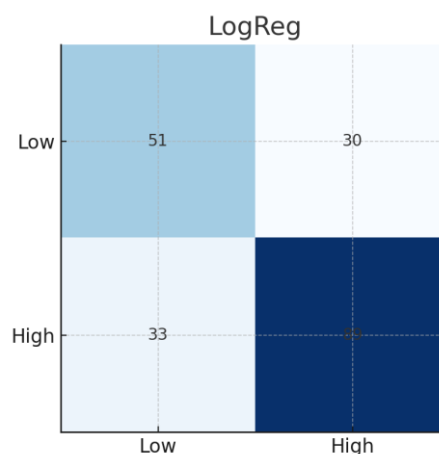
Grafik perbandingan akurasi ditampilkan pada Gambar 3.



Gambar 3. Perbandingan Akurasi

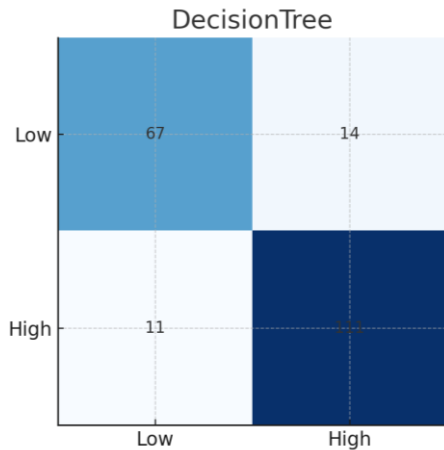
Gambar 3 menunjukkan bahwa model *Decision Tree* memiliki akurasi tertinggi (87,68%), diikuti oleh *Random Forest* (87,19%). *Logistic Regression* memiliki performa paling rendah (68,97%), disusul *SVM* (75,37%) dan *KNN* (77,34%).

Selain akurasi, *confusion matrix* digunakan untuk menilai kesalahan klasifikasi masing-masing model. Visualisasi *confusion matrix* setiap model dan interpretasi *confusion matrix* dijelaskan sebagai berikut,



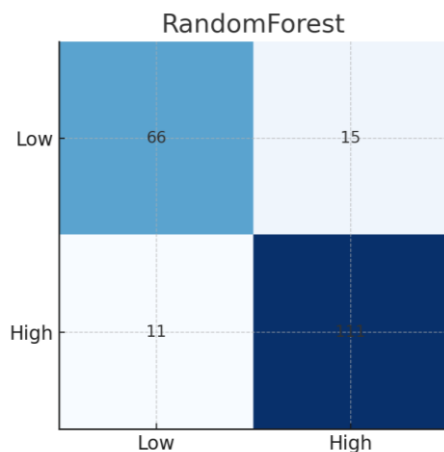
Gambar 4. Visualisasi *confusion matrix* model *Logistic Regression*

Berdasarkan Gambar 4 model *Logistic Regression* dengan tingkat akurasi 68,97 % dan benar *low risk*: 51, salah *low risk* → *high risk*: 30, benar *high risk*: 89, salah *high risk* → *low risk*: 33. model ini banyak salah memprediksi *high risk* menjadi *low risk* (33 kasus), sehingga kurang cocok untuk kasus medis yang membutuhkan *recall* tinggi.



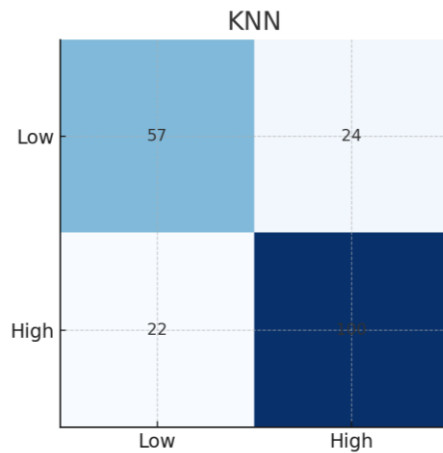
Gambar 5. Visualisasi *confusion matrix* model *Decision Tree*

Berdasarkan Gambar 5 model *Decision Tree* dengan tingkat akurasi 87,68 %, hampir seluruh data *high risk* diprediksi benar (111/122). Kesalahan model rendah pada kedua kelas serta menjadi model terbaik dalam penelitian.



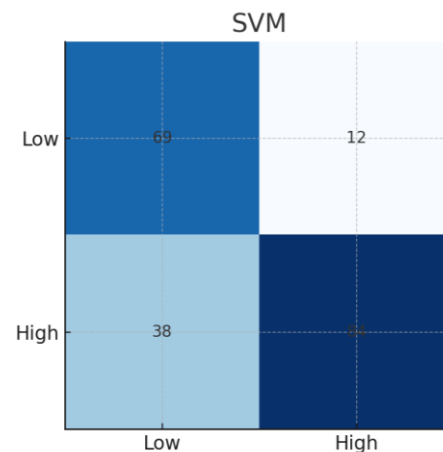
Gambar 6. Visualisasi *confusion matrix* model *Random Forest*

Berdasarkan Gambar 6 model *Random Forest* dengan tingkat akurasi 87,19 %, mirip dengan *Decision Tree*, namun sedikit lebih banyak memprediksi *low risk* secara salah serta lebih stabil dan tidak terlalu *overfitting*.



Gambar 7. Visualisasi *confusion matrix* model *KNN*

Berdasarkan Gambar 7 model *KNN* dengan tingkat akurasi 77,34%, kesalahan prediksi terjadi pada kedua kelas, terutama *low risk* → *high risk* serta sensitif terhadap skala dan *noise data*.



Gambar 8. Visualisasi *confusion matrix* model *SVM*

Berdasarkan Gambar 8 model *SVM* dengan tingkat akurasi 75,37%, *recall high risk* rendah (68,85%), artinya banyak ibu hamil berisiko tinggi yang salah diklasifikasi sebagai risiko rendah serta tidak cukup kuat untuk klasifikasi medis.

Perhitungan manual akurasi, *precision*, *recall*, dan *f1-score* untuk model *decision tree* dengan *confusion matrix* yang ditunjukkan pada Gambar 5 sebagai berikut

1. Accuracy

$$Accuracy = \frac{TP + TN}{Total} = \frac{111 + 67}{203} = \frac{178}{203} = 0,876$$

2. *Precision High Risk* (Positif)

$$Precision = \frac{TP}{TP + FP} = \frac{111}{111 + 14} = \frac{111}{125} = 0,888$$

3. *Recall High Risk* (Positif)

$$Recall = \frac{TP}{TP + FN} = \frac{111}{111 + 11} = \frac{111}{122} = 0,9098$$

4. *F1-Score*

$$F1 = 2x \frac{precision \times recall}{precision + recall}$$

$$F1 = 2x \frac{0,888 \times 0,9098}{0,888 + 0,9098} = 2x \frac{0,809}{1,7978} = 0,8988$$

Berdasarkan hasil, Model *Decision Tree* adalah model yang memberikan performa paling baik dengan akurasi 87,68%, disusul Random Forest dengan akurasi 87,19%. Kedua model ini efektif dalam menangkap pola non-linear pada data klinis ibu hamil, terutama interaksi antara tekanan darah sistolik, tekanan darah diastolik, dan kadar gula darah, yang merupakan parameter dominan dalam menentukan risiko preeklampsia.

Model berbasis pohon bekerja lebih baik karena

1. Mampu melakukan pemisahan data secara hierarkis berdasarkan fitur paling penting.
2. Tidak memerlukan normalisasi data
3. Dapat menangani interaksi kompleks antar variabel kesehatan
4. Tidak sensitif terhadap *outlier*

Sementara itu, *Logistic Regression* bekerja kurang optimal karena model linear tidak dapat merepresentasikan pola hubungan non-linear pada dataset ini. Hal yang sama terjadi pada *SVM* dan *KNN* yang masih kalah performa dibandingkan model berbasis pohon.

Performa kedua model terbaik (DT dan RF) menunjukkan bahwa metode *decision tree* sangat cocok untuk diterapkan dalam penelitian deteksi risiko preeklampsia. Dengan akurasi di atas 87%, model ini sudah sangat berpotensi diterapkan sebagai sistem bantu diagnosis di fasilitas pelayanan kesehatan.

IV. KESIMPULAN

Pengujian menggunakan 30% data testing dan 70% data training dengan enam algoritma: *decision tree*,

naive bayes, *SVM*, *KNN*, *random forest*, dan *logistic regression*. Hasil pengujian menunjukkan akurasi algoritma sebagai berikut: *decision tree* 99%, *naive bayes* 48%, *SVM* 95%, *KNN* 99%, *random forest* 99%, dan *logistic regression* 74%. Oleh karena itu, algoritma *decision tree*, *KNN*, dan *random forest* dengan akurasi 99% lebih efektif untuk mendeteksi stunting pada balita.

V. UCAPAN TERIMA KASIH

Terima kasih atas dukungan dana penelitian yang diberikan oleh Politeknik Kaltara melalui LPPM.

VI. REFERENSI

- Aisha, P., Fathurahman, M., & Prangga, S. (2024). Implementasi Metode Neighbor Weighted K-Nearest Neighbor Pada Pengklasifikasian Status Gizi Balita Di Wilayah Kerja Puskesmas Wonorejo Kota Samarinda. *VARIANCE: Journal Of Statistics And Its Applications*, 6(1), 11–20. <https://doi.org/10.30598/Variancevol6iss1page11-20>
- Chazar, C., & Erawan, B. (2020). Machine Learning Diagnosis Kanker Payudara Menggunakan Algoritma Support Vector Machine. *INFORMASI (Jurnal Informatika Dan Sistem Informasi)*, 12(1), 67–80. <https://doi.org/10.37424/Informasi.V12i1.48>
- Fauzan Adzim, Budianita, E., Nazir, A., & Syafria, F. (2023). Klasifikasi Status Stunting Balita Menggunakan Metode C4.5 Berbasis Web. *Zonasi: Jurnal Sistem Informasi*, 5(3), 215–225. <https://doi.org/10.31849/Zn.V5i3.15828>
- Fitrilina, F., Albbi, M., Agustian, I., Herawati, A., & Massardi, N. A. (2021). Sistem Peringatan Awal Resiko Preeklampsia Pada Kehamilan Menggunakan Metoda Certainty Factor Dan Android. *Jurnal Nasional Teknik Elektro*, 10(1), 45. <https://doi.org/10.25077/jnte.V10n1.896.2021>
- Hanif, K. H., Muntiari, N. R., & Ramadhani, P. A. (2022). Penerapan Metode Certainty Factor Untuk Mendiagnosa Penyakit Preeklampsia Pada Ibu Hamil Dengan Menggunakan Bahasa Pemrograman Python. *Insect (Informatics And Security): Jurnal Teknik Informatika*, 7(2), 63–71.
- Kazerouni, F., Bayani, A., Asadi, F., Saeidi, L., Parvizi, N., & Mansoori, Z. (2020). Type2 Diabetes Mellitus Prediction Using Data Mining Algorithms Based On The Long-Noncoding Rnas Expression: A Comparison Of Four Data

- Mining Approaches. *BMC Bioinformatics*, 21(1), 1-13.
<https://doi.org/10.1186/S12859-020-03719-8>
- Makmur Mulyono¹, Elvia Budianita², Alwis Nazir³, F. S. (2023). Klasifikasi Status Stunting Balita Menggunakan Metode Naïve Bayes Gaussian Berbasis Web. *Zonasi: Jurnal Sistem Informasi*, 5(3), 215-225.
<https://doi.org/10.31849/Zn.V5i3.15828>
- Muntiari, N. R., & Hanif, K. H. (2022). Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma Machine Learning. *Jurnal Ilmu Komputer Dan Teknologi*, 3(1), 1-6.
<https://doi.org/10.35960/ikomti.V3i1.766>
- Muntiari, N. R., Hanif, K. H., & Nisa, I. C. (2021). Perbandingan Algoritma Regresi Logistik , Support Vector Machine , Dan Gradient Boosting Pada Analisis Sentimen Data Komentar Siswa. 2(1), 1-7.
- Muntiari, N. R., Nisa, K., Sandi, A. S., Ashari, I. A., Kharis Hudaiby Hanif, A., & Dwinanto, R. W. (2023). *Comparison Of Random Forest Algorithm, Support Vector Machine, And K-Nearest Neighbor For Diabetes Disease Classification*. May.
- Nour, M., & Polat, K. (2020). Automatic Classification Of Hypertension Types Based On Personal Features By Machine Learning Algorithms. *Mathematical Problems In Engineering*, 2020, 1-14.
<https://doi.org/10.1155/2020/2742781>
- Nurmalasari, M. D., Kusri, K., & Sudarmawan, S. (2021). Komparasi Algoritma Naive Bayes Dan K-Nearest Neighbor Untuk Membangun Pengetahuan Diagnosa Penyakit Diabetes. *Jurnal Komtika (Komputasi Dan Informatika)*, 5(1), 52-59.
<https://doi.org/10.31603/Komtika.V5i1.5140>
- Oktavianto, H., & Handri, R. P. (2020). Analisis Klasifikasi Kanker Payudara Menggunakan Algoritma Naive Bayes. *INFORMAL: Informatics Journal*, 4(3), 117.
<https://doi.org/10.19184/Isj.V4i3.14170>
- Priya, R., & Abbas Rizvi, S. W. (2020). SVM & Ga-Clustering Based Feature Selection Approach For Breast Cancer Detection. *International Journal On Soft Computing, Artificial Intelligence And Applications*, 1(10), 1-10.
<https://doi.org/10.5121/Ijssci.2020.9401>
- Rahayu, W. I., Prianto, C., & Novia, E. A. (2021). Perbandingan Algoritma K-Means Dan Naive Bayes Untuk Memprediksi Prioritas Pembayaran Tagihan Rumah Sakit Berdasarkan Tingkat Kepentingan Pada PT. Pertamina (Persero). *Jurnal Teknik Informatika*, 13(2), 1-8.
- Ramadhani, Ramadhanu, & Taufik Hidayat. (2024). Metode Machine Learning Untuk Klasifikasi Data Gizi Balita Dengan Algoritma Naive Bayes, KNN Dan Decision Tree. *Jurnal SIMETRIS*, 15(1), 57-68.
- Rawansyah, Lestari, V. A., & Anita, S. (2020). Ibu Hamil Menggunakan Metode Fuzzy Logic Dan Certainty Factor. *Seminar Informatika Aplikatif Polinema*, 221-225.
<http://jurnali.polinema.ac.id/index.php/SIAP/article/view/773>
- Taye, M. M. (2023). Understanding Of Machine Learning With Deep Learning: *Computers MDPI*, 12(91), 1-26.
- Tjahjadi, H., & Ramli, K. (2020). Noninvasive Blood Pressure Classification Based On Photoplethysmography Using K-Nearest Neighbors Algorithm: A Feasibility Study. *Information (Switzerland)*, 11(2).
<https://doi.org/10.3390/info11020093>
- Wiratama, Y., & Aziz, R. Z. A. (2024). *Perbandingan Prediksi Penyakit Stunting Balita Menggunakan Algoritma Support Vektor Machine Dan Random Forest*. 6(2).
<https://doi.org/10.47065/Bits.V6i2.5543>
- Yudha Prawira, T., Fadlil, A., & Dahlan Jalan Soepono, A. (N.D.). Sistem Pendukung Keputusan / Decision Support System Untuk Memprediksi Stok Barang Pada Retail Mini Market Berbasis Mobile Dengan Pendekatan Analisis Menggunakan Metode Algoritma Apriori. In *Symposium Nasional Teknologi Terapan (SNTT)* (Vol. 5).