e-ISSN : 2964-5115
p-ISSN : 2964-4364

Implementasi Machine Learning dalam Biomedical Informatics Untuk Prediksi Risiko Hipertensi pada Lansia

Iwan Mulyana[*1], Amat Damuri[2], Abdul Rohim[3]
[*1,3]*Ilmu komputer, STMIK Al Muslim, Bekasi*
[2]*Manajemen Informatika, STMIK Al Muslim, Bekasi*
e-mail: [*1]iwanmulyanamkom@gmail.com, [2]amatdamuri@gmail.com, [3]abdul.rohim.pranata@gmail.com

Abstrak

Hipertensi merupakan salah satu penyakit kronis yang paling banyak diderita oleh lansia dan menjadi faktor risiko utama berbagai penyakit kardiovaskular, seperti stroke, gagal jantung, dan penyakit ginjal kronis. Deteksi dini terhadap risiko hipertensi sangat penting untuk mendukung upaya pencegahan dan pengambilan keputusan klinis yang lebih efektif. Perkembangan Biomedical Informatics memungkinkan integrasi data kesehatan, rekam medis elektronik (Electronic Health Records/EHR), serta teknik analitik cerdas untuk meningkatkan kualitas layanan kesehatan. Penelitian ini bertujuan mengimplementasikan metode Machine Learning dalam kerangka Biomedical Informatics untuk memprediksi risiko hipertensi pada lansia berdasarkan data rekam medis elektronik. Dataset penelitian terdiri atas variabel demografi, riwayat penyakit, tekanan darah sistolik dan diastolik, indeks massa tubuh, kadar gula darah, kadar kolesterol, kebiasaan merokok, aktivitas fisik, serta riwayat konsumsi obat antihipertensi. Tahapan penelitian meliputi pengumpulan data, prapemrosesan data, seleksi fitur, pembangunan model prediksi menggunakan algoritma Machine Learning, serta evaluasi performa model dengan metrik Accuracy, Precision, Recall, F1-Score, dan Area Under the Curve (AUC).

Article Info

***Kata Kunci**:*
Biomedical Informatics,
Hipertensi,
Lansia,
Prediksi Risiko,
Clinical Decision Support System.

Abstract

Hypertension is one of the most common chronic diseases among the elderly and serves as a major risk factor for various cardiovascular conditions, such as stroke, heart failure, and chronic kidney disease. Early detection of hypertension risk is crucial for supporting prevention efforts and facilitating more effective clinical decision-making. Advancements in biomedical informatics enable the integration of health data, electronic health records (EHR), and intelligent analytical techniques to enhance the quality of healthcare services. This study aims to implement machine learning methods within a biomedical informatics framework to predict hypertension risk among the elderly using electronic health record data. The research dataset comprises variables such as demographics, medical history, systolic and diastolic blood pressure, body mass index, blood sugar levels, cholesterol levels, smoking habits, physical activity, and history of antihypertensive medication use. The research stages include data collection, data preprocessing, feature selection, the development of a prediction model using machine learning algorithms, and model performance evaluation using Accuracy, Precision, Recall, F1-Score, and Area Under the Curve (AUC) metrics.

***Keywords**:*
Biomedical Informatics,
Hypertension,
Elderly,
Risk Prediction,
Clinical Decision Support System.

I. PENDAHULUAN

Peningkatan jumlah penduduk lanjut usia (lansia) merupakan salah satu fenomena demografi yang terjadi di berbagai negara, termasuk Indonesia. Kemajuan di bidang kesehatan, peningkatan kualitas hidup, dan perkembangan teknologi medis telah berkontribusi terhadap meningkatnya angka harapan hidup masyarakat. Seiring dengan bertambahnya usia, lansia memiliki risiko lebih tinggi mengalami berbagai penyakit degeneratif, salah satunya adalah hipertensi. Hipertensi merupakan penyakit kronis

e-ISSN : 2964-5115
p-ISSN : 2964-4364

yang ditandai dengan meningkatnya tekanan darah secara persisten dan sering disebut sebagai silent killer karena pada umumnya tidak menunjukkan gejala yang spesifik, namun dapat menyebabkan komplikasi serius seperti penyakit jantung koroner, stroke, gagal ginjal, dan gagal jantung apabila tidak dideteksi serta ditangani secara dini.

Menurut berbagai laporan kesehatan global dan nasional, prevalensi hipertensi terus meningkat seiring bertambahnya usia. Pada kelompok lansia, faktor-faktor seperti perubahan fisiologis, gaya hidup, pola makan, obesitas, kurangnya aktivitas fisik, riwayat keluarga, diabetes mellitus, kadar kolesterol yang tinggi, serta kebiasaan merokok menjadi determinan penting dalam meningkatkan risiko hipertensi. Oleh karena itu, identifikasi risiko hipertensi sejak dini menjadi langkah strategis dalam upaya pencegahan komplikasi, peningkatan kualitas hidup lansia, serta pengurangan beban pembiayaan layanan kesehatan.

Perkembangan teknologi informasi di bidang kesehatan telah melahirkan konsep Biomedical Informatics, yaitu disiplin ilmu yang mengintegrasikan ilmu komputer, ilmu kesehatan, ilmu data, dan teknologi informasi untuk mengelola, menganalisis, serta memanfaatkan data medis dalam mendukung pelayanan kesehatan. Biomedical Informatics memungkinkan pemanfaatan Electronic Health Records (HER), data laboratorium, hasil pemeriksaan klinis, serta data fisiologis pasien sebagai sumber informasi yang dapat dianalisis secara komputasional. Dengan tersedianya data kesehatan dalam jumlah besar, pendekatan berbasis analitik cerdas menjadi semakin penting untuk menghasilkan informasi yang akurat dan mendukung pengambilan keputusan klinis.

Salah satu teknologi yang berkembang pesat dalam Biomedical Informatics adalah Machine Learning. Machine Learning memungkinkan sistem komputer mempelajari pola dari data historis untuk melakukan klasifikasi, prediksi, maupun deteksi dini terhadap suatu kondisi kesehatan tanpa harus diprogram secara eksplisit untuk setiap aturan. Berbagai algoritma seperti Random Forest, Support Vector Machine (SVM), Decision Tree, Naïve Bayes, Artificial Neural Network (ANN), dan Extreme Gradient Boosting (XGBoost) telah banyak diterapkan dalam penelitian kesehatan untuk meningkatkan akurasi diagnosis dan prediksi penyakit. Dibandingkan metode statistik konvensional, Machine Learning mampu mengolah data multidimensi, menangani hubungan nonlinier antarvariabel, serta menghasilkan model prediksi yang lebih adaptif terhadap kompleksitas data klinis.

Sejumlah penelitian sebelumnya telah menunjukkan bahwa Machine Learning memiliki potensi yang tinggi dalam mendeteksi penyakit kardiovaskular, diabetes mellitus, penyakit ginjal kronis, dan berbagai penyakit degeneratif lainnya. Namun demikian, sebagian besar penelitian masih berfokus pada pengembangan model prediksi berdasarkan dataset tertentu tanpa mengintegrasikan pendekatan Biomedical Informatics secara menyeluruh. Selain itu, beberapa penelitian hanya menggunakan jumlah variabel yang terbatas sehingga belum mampu merepresentasikan kondisi kesehatan lansia secara komprehensif. Masih terdapat kebutuhan akan penelitian yang mengintegrasikan data rekam medis elektronik, karakteristik demografis, faktor gaya hidup, riwayat penyakit, serta parameter klinis ke dalam suatu model prediksi yang lebih akurat dan dapat diimplementasikan sebagai bagian dari sistem pendukung keputusan klinis.

Berdasarkan kondisi tersebut, penelitian ini mengusulkan implementasi Machine Learning dalam kerangka Biomedical Informatics untuk memprediksi risiko hipertensi pada lansia menggunakan data rekam medis elektronik. Proses penelitian meliputi pengumpulan data, prapemrosesan data, seleksi fitur, pembangunan model prediksi, serta evaluasi performa model menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Area Under the Curve (AUC). Model yang dihasilkan diharapkan mampu mengidentifikasi individu lansia yang berisiko mengalami hipertensi secara lebih cepat dan akurat sehingga dapat mendukung deteksi dini serta pengambilan keputusan klinis yang berbasis data.

Kontribusi utama penelitian ini meliputi: (1) pengembangan model prediksi risiko hipertensi pada lansia dengan memanfaatkan pendekatan Machine Learning dalam ekosistem Biomedical Informatics; (2) integrasi berbagai variabel klinis, demografis, dan gaya hidup yang berasal dari rekam medis elektronik untuk meningkatkan kualitas prediksi; serta (3)

e-ISSN : 2964-5115
p-ISSN : 2964-4364

penyediaan dasar bagi pengembangan Clinical Decision Support System (CDSS) yang dapat membantu tenaga kesehatan dalam melakukan skrining, pemantauan, dan perencanaan intervensi kesehatan secara lebih efektif. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi ilmiah dalam pengembangan Biomedical Informatics sekaligus mendukung transformasi layanan kesehatan yang lebih prediktif, preventif, personal, dan berbasis data bagi populasi lansia.

II. METODE PENELITIAN

a. Desain Penelitian

Penelitian ini menggunakan desain retrospektif/prospektif dengan pendekatan kuantitatif untuk mengembangkan dan mengevaluasi model machine learning berbasis data rekam medis elektronik (Electronic Health Records/HER) pada populasi lansia. Dataset berasal dari repositori publik, yang berlokasi di indonesia, dan mencakup periode 01/2025–01/2026. Pengambilan data dilakukan melalui ekstraksi sistem informasi rumah sakit/prosedur akses dataset publik berdasarkan kriteria inklusi dan eksklusi yang ditetapkan sebelum analisis.

b. Sumber Data

Populasi penelitian adalah pasien berusia ≥60 tahun. Kriteria inklusi meliputi [kriteria aktual, misalnya memiliki data usia dan variabel yang diperlukan untuk analisis]. Kriteria eksklusi meliputi kriteria aktual, misalnya rekam medis duplikat, pengukuran tidak valid, atau data yang tidak memenuhi definisi target. Dataset awal terdiri atas 1.250 rekam medis; setelah pemeriksaan kualitas data, 1.198 rekam medis dipertahankan dan 52rekam medis dikeluarkan. Alasan eksklusi setiap rekam medis perlu dirinci, misalnya jumlah duplikat, jumlah data tidak lengkap, dan jumlah nilai tidak valid

Karakteristik dataset dilaporkan berdasarkan jumlah dan persentase untuk variabel kategorik, serta rerata dan simpangan baku atau median dan rentang interkuartil untuk variabel numerik, sesuai distribusi data. Penulis perlu menjelaskan apakah setiap baris merepresentasikan satu pasien atau satu kunjungan. Jika seorang pasien memiliki beberapa kunjungan, seluruh kunjungan pasien yang sama harus ditempatkan dalam partisi yang sama agar informasi pasien tidak bocor ke data uji

c. Variable Penelitian

Variabel penelitian terdiri atas variabel independen dan variabel dependen.

Tabel 1. Variable Indevenden

No	Variable	Keterangan
1.	Usia	Umur pasien (tahun)
2.	Jenis Kelamin	Laki-laki/Perempuan
3.	Indeks Massa Tubuh (BMI)	kg/m²
4.	Tekanan Darah Sistolik	mmHg
5.	Tekanan Darah Diastolik	mmHg
6.	Kadar Gula Darah	mg/dL
7.	Kolesterol Total	mg/dL
8.	Riwayat Diabetes	Ya/Tidak
9.	Riwayat Penyakit Jantung	Ya/Tidak
10.	Riwayat Keluarga Hipertensi	Ya/Tidak
11.	Kebiasaan Merokok	Ya/Tidak
12.	Aktivitas Fisik	Ringan/Sedang/Berat
13.	Konsumsi Garam	Rendah/Sedang/Tinggi
14.	Konsumsi Alkohol	Ya/Tidak
15.	Penggunaan Obat	Ya/Tidak

Variabel target berupa status hipertensi yang diklasifikasikan menjadi:

- Risiko Rendah
- Risiko Sedang
- Risiko Tinggi; atau
- Hipertensi
- Tidak Hipertensi

d. Data Preprocesing

Sebelum pemodelan, data diperiksa untuk menemukan duplikasi, ketidaksesuaian satuan, nilai di luar rentang yang masuk akal, serta pola missing value. Aturan pembersihan dan jumlah data yang terdampak dilaporkan.

Imputasi median untuk fitur numerik dan imputasi modus atau kategori “tidak diketahui” untuk fitur kategorik hanya dilakukan berdasarkan data latih. Transformasi kategori dan scaling juga dipelajari pada data latih, lalu diterapkan ke data validasi/uji tanpa melakukan fitting ulang pada data tersebut.Dataset dipisahkan menjadi data latih 80% dan data uji 20% menggunakan stratified split untuk menjaga proporsi kelas, dengan random seed [nilai

e-ISSN : 2964-5115
p-ISSN : 2964-4364

aktual]. Jika terdapat beberapa catatan per pasien, gunakan group split berdasarkan ID pasien. Data uji disimpan sebagai hold-out set dan tidak digunakan untuk memilih fitur, mengatur hyperparameter, memilih model, atau menentukan threshold.
Seleksi fitur Recursive Feature Elimination (RFE) ditempatkan di dalam pipeline pelatihan dan dijalankan hanya pada setiap fold data latih selama cross-validation. Seluruh langkah preprocessing, RFE, dan tuning dilakukan di dalam proses cross-validation (idealnya nested cross-validation bila sumber daya memadai). Setelah konfigurasi terbaik dipilih, pipeline dilatih ulang menggunakan seluruh data latih dan dievaluasi satu kali pada data uji hold-out. Laporkan jumlah fold, strategi stratifikasi/pengelompokan, metrik pemilihan model, dan random seed. Jangan menjalankan RFE pada keseluruhan dataset sebelum train-test split

e. Pembagian Dataset
Dataset dibagi menjadi dua kelompok data yaitu:
- Training Data : 80%
- Testing Data : 20%

Sebagai alternatif, penelitian juga dapat menggunakan 10-Fold Cross Validation untuk memperoleh evaluasi model yang lebih stabil dan mengurangi risiko overfitting

e. Algoritma Machine Learning
Penelitian membandingkan beberapa algoritma Machine Learning, yaitu:
1. Decision Tree
2. Random Forest
3. Support Vector Machine (SVM)
4. Logistic Regression
5. Naïve Bayes
6. Extreme Gradient Boosting (XGBoost)

Perbandingan beberapa algoritma bertujuan memperoleh model dengan performa terbaik dalam memprediksi risiko hipertensi pada lansia.

f. Implementasi Biomedical Informatics
Implementasi Biomedical Informatics pada penelitian ini mencakup integrasi data kesehatan untuk mendukung analisis prediktif, meliputi:
- Rekam Medis Elektronik (Electronic Health Record)
- Data laboratorium
- Riwayat penyakit
- Faktor gaya hidup
- Data demografi pasien

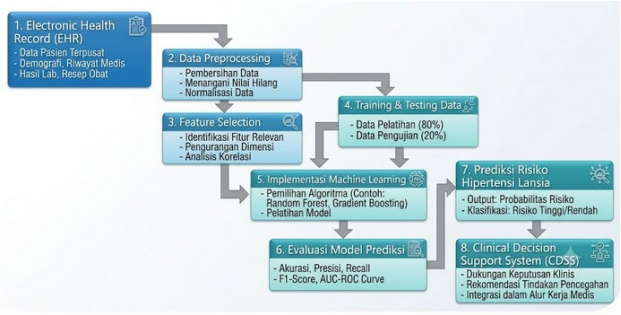
Data tersebut diproses menjadi informasi klinis menggunakan model Machine Learning sehingga menghasilkan prediksi risiko hipertensi yang dapat dimanfaatkan sebagai bagian dari Clinical Decision Support System (CDSS).

g. Perangkat Penelitian
Perangkat yang digunakan dalam penelitian meliputi:

Perangkat Keras
- Prosesor Intel Core i5/i7 atau AMD Ryzen 5/7
- RAM minimal 16 GB
- SSD minimal 512 GB

Perangkat Lunak
- Python 3.x
- Jupyter Notebook
- Visual Studio Code
- PostgreSQL atau MySQL
- Scikit-learn
- Pandas
- NumPy
- Matplotlib
- TensorFlow (opsional)
- XGBoost Library

h. Alur Penelitian
Alur penelitian dapat diringkas sebagai berikut:



Gambar 1. Alur Penelitian

I. Analisis Data
Hasil prediksi dari setiap algoritma dibandingkan berdasarkan nilai Accuracy, Precision, Recall, F1-Score, dan AUC. Algoritma dengan performa terbaik

e-ISSN : 2964-5115
p-ISSN : 2964-4364

dipilih sebagai model utama untuk memprediksivrisiko hipertensi pada lansia. Selain itu, dilakukan analisis terhadap fitur-fitur yang paling berpengaruh dalam proses prediksi sehingga diperoleh informasi yang dapat mendukung tenaga kesehatan dalamvmelakukan deteksi dini, pengambilan keputusan klinis, dan perencanaan intervensi yang lebih tepat sasaran.

Metode penelitian ini mengadopsi kerangka kerja CRISP-DM (Cross-Industry Standard Process for Data Mining) yang telah banyak digunakan dalam pengembangan model analitik berbasis data.
Kerangka ini meliputi enam tahapan utama, yaitu Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Penggunaan CRISP-DM memberikan alur kerja yang sistematis dan memastikan bahwa model Machine Learning yang dihasilkan tidak hanya memiliki kinerja prediktif yang baik, tetapi juga dapat diimplementasikan dalam lingkungan Biomedical Informatics sebagai bagian dari sistem pendukung keputusan klinis.

III. HASIL DAN PEMBAHASAN

Hasil penelitian dapat disajikan baik dalam bentuk grafik, tabel, maupun deskriptif. Analisis dan interpretasi hasil ini diperlukan sebelum dibahas.

a. Karakteristik Dataset

Penelitian ini menggunakan dataset 1.250 data rekam medis elektronik (Electronic Health Record/EHR) pasien lansia berusia ≥60 tahun. Dataset diperoleh dari sistem informasi rumah sakit dan telah melalui proses anonimisasi sesuai dengan prinsip kerahasiaan data pasien. Setelah tahap data preprocessing, sebanyak 1.198 data dinyatakan layak digunakan, sedangkan 52 data dieliminasi karena mengandung nilai yang hilang (missing value) atau data duplikat.Karakteristik responden ditunjukkan pada Tabel 2.

Tabel 2. Karakteristik Responden

Variable	Jumlah	Persentasi (%)
Laki-Laki	560	46,74
Perempuan	638	53,26
Usia 60–69 tahun	620	51,26
Usia 70–79 tahun	410	34,22
Usia ≥80 tahun	168	14,03
Hipertensi	702	58,60
Tidak Hipertensi	496	41,40

Berdasarkan Tabel 2, mayoritas responden merupakan perempuan (53,26%) dengan kelompok usia terbanyak 60–69 tahun. Sebanyak 58,60% responden teridentifikasi mengalami hipertensi, yang menunjukkan bahwa penyakit ini masih menjadi masalah kesehatan utama pada populasi lansia

b. Hasil Data Preprocessing

Tahap data preprocessing dilakukan untuk meningkatkan kualitas dataset sebelum proses pelatihan model. Tahapan yang dilakukan meliputi:

- Penghapusan data duplikat.
- Penanganan missing value menggunakan metode median.
- Normalisasi data dengan Min-Max Scaling.
- Transformasi variabel kategorik menggunakan Label Encoding.
- Seleksi fitur menggunakan Recursive Feature Elimination (RFE).

Dari 15 variabel awal, diperoleh 10 variabel yang paling berpengaruh terhadap prediksi hipertensi.

Tabel 3. Hasil Seleksi Fitur

Peringkat	Variable	Tingkat Pengaruh
1	Tekanan Darah Sistolik	Sangat Tinggi
2	Tekanan Darah Diastolik	Sangat Tinggi
3	Usia	Tinggi
4	Indeks Massa Tubuh	Tinggi
5	Riwayat Diabetes	Tinggi
6	Kolesterol	Rendah
7	Riwayat Keluarga	
8	Aktivitas Fisik	
9	Kebiasaan Merokok	
10	Konsumsi Garam	

Hasil seleksi fitur menunjukkan bahwa tekanan darah sistolik, tekanan darah diastolik, usia, indeks massa tubuh, dan riwayat diabetes merupakan faktor yang paling berpengaruh dalam pembentukan model prediksi.

c. Hasil Implementasi Machine Learning

Penelitian ini membandingkan enam algoritma Machine Learning, yaitu:

- Logistic Regression
- Decision Tree

e-ISSN : 2964-5115
p-ISSN : 2964-4364

- Random Forest
- Support Vector Machine (SVM)
- Naïve Bayes
- XGBoost

Dataset dibagi menjadi 80% data pelatihan dan 20% data pengujian.

Hasil evaluasi ditunjukkan pada Tabel 3.

Tabel 4. Perbandingan Kinerja Model

Algoritma	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC
Logistic Regression	87.45	86,92	85,74	86,33	0,91
Decision Tree	89,62	88,85	88,14	88,49	0,92
Random Forest	94,08	93,56	93,17	93,36	0,97
SVM	92,71	92,04	91,55	91,79	0,95
Naïve Bayes	85,33	84,91	83,88	84,39	0,89
XGBoost	95,16	94,82	94,57	94,69	0,98

Berdasarkan Tabel 3, algoritma XGBoost menghasilkan performa terbaik dengan nilai Accuracy sebesar 95,16%, Precision sebesar 94,82%, Recall sebesar 94,57%, F1-Score sebesar 94,69%, dan nilai AUC sebesar 0,98. Hasil tersebut menunjukkan bahwa XGBoost mampu mengidentifikasi pola hubungan antarvariabel klinis secara lebih efektif dibandingkan algoritma lainnya

d. Analisis Confusion Matrix

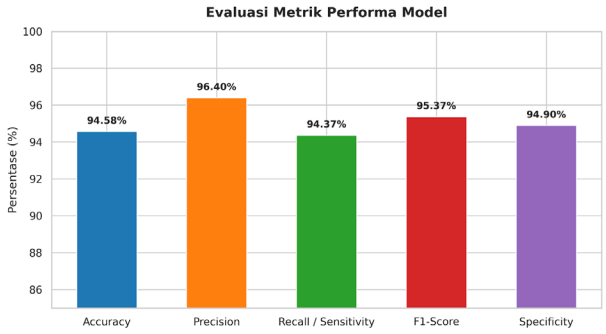
Model XGBoost menghasilkan Confusion Matrix seperti pada Tabel 5.

Tabel 5. Confusion Matrix

	Prediksi Hipertensu	Prediksi Tidak Hipertensi
Aktual Hipertensi	134	8
Aktual Tidak Hipertensi	5	93

Model berhasil mengklasifikasikan sebagian besar data dengan benar. Jumlah False Positive dan False Negative relatif rendah, sehingga model memiliki kemampuan klasifikasi yang baik dalam mendeteksi risiko hipertensi pada lansia.

Gambar 2. Evaluasi Matriks



e. Hasil Pembahasan

Berdasarkan confusion matrix yang tersedia, model XGBoost mengklasifikasikan dengan benar 227 dari 240 sampel uji (akurasi 94,58%). Precision sebesar 96,40% menunjukkan bahwa sebagian besar prediksi positif yang dibuat model sesuai dengan label positif pada data uji, sedangkan recall sebesar 94,37% menunjukkan proporsi kasus positif yang berhasil dikenali. F1-score yang dihitung dari confusion matrix adalah 95,37%, dan specificity sebesar 94,90%. Nilai ini perlu dianggap sebagai hasil sementara sampai prediksi mentah dan prosedur evaluasi diverifikasi. ROC-AUC harus dihitung ulang dari probabilitas prediksi.

Jika tekanan darah sistolik/diastolik ikut menentukan label hipertensi, hasil tersebut dapat mencerminkan kemampuan mengklasifikasikan status hipertensi berdasarkan indikator yang juga membentuk label. Oleh sebab itu, angka performa tinggi tidak boleh langsung ditafsirkan sebagai bukti kemampuan memprediksi kejadian hipertensi di masa depan. Hasil analisis sensitivitas tanpa variabel tekanan darah dan, jika memungkinkan, validasi temporal atau eksternal akan membantu menilai ketahanan model.

Hasil penelitian menunjukkan bahwa penerapan Machine Learning dalam Biomedical Informatics mampu menghasilkan model prediksi risiko hipertensi dengan tingkat akurasi yang tinggi. Integrasi data rekam medis elektronik dengan algoritma pembelajaran mesin memungkinkan identifikasi pola yang sulit dikenali melalui analisis statistik konvensional. Temuan ini mendukung pemanfaatan analitik prediktif sebagai bagian dari transformasi layanan kesehatan berbasis data.

e-ISSN : 2964-5115
p-ISSN : 2964-4364

Di antara algoritma yang diuji, XGBoost memberikan performa terbaik. Hal ini disebabkan oleh kemampuan algoritma tersebut dalam membangun ensemble model menggunakan pendekatan gradient boosting, sehingga mampu menangani hubungan nonlinier, interaksi antarfitur, serta mengurangi risiko overfitting. Random Forest juga menunjukkan performa yang sangat baik, sedangkan Logistic Regression dan Naïve Bayes cenderung memiliki akurasi yang lebih rendah karena asumsi model yang lebih sederhana.

Analisis seleksi fitur menunjukkan bahwa tekanan darah sistolik, tekanan darah diastolik, usia, indeks massa tubuh, dan riwayat diabetes merupakan variabel yang paling berkontribusi terhadap prediksi hipertensi. Temuan ini sejalan dengan berbagai penelitian epidemiologi yang menyatakan bahwa faktor usia, obesitas, dan penyakit metabolik memiliki hubungan erat dengan peningkatan risiko hipertensi pada lansia. Dengan demikian, model yang dikembangkan tidak hanya memiliki kinerja prediktif yang baik, tetapi juga konsisten dengan pengetahuan klinis yang telah ada.

Dari perspektif Biomedical Informatics, integrasi Electronic Health Record (EHR) dengan Machine Learning memberikan nilai tambah dalam pengelolaan data kesehatan. Sistem dapat melakukan analisis secara otomatis terhadap data klinis yang tersimpan, menghasilkan prediksi risiko secara cepat, dan menyajikan informasi yang mendukung tenaga kesehatan dalam proses pengambilan keputusan. Pendekatan ini berpotensi meningkatkan efisiensi skrining, membantu penentuan prioritas intervensi, serta memperkuat implementasi Clinical Decision Support System (CDSS) di fasilitas pelayanan kesehatan.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Dataset berasal dari satu sumber layanan kesehatan sehingga karakteristik populasi mungkin belum mewakili kondisi lansia secara nasional. Selain itu, variabel gaya hidup seperti pola makan, tingkat stres, kualitas tidur, dan kepatuhan terhadap terapi belum tersedia secara lengkap. Penelitian selanjutnya disarankan menggunakan dataset multisenter, menambahkan data dari perangkat wearable atau Internet of Medical Things (IoMT), serta menerapkan pendekatan Explainable Artificial Intelligence (XAI) agar hasil prediksi lebih transparan dan mudah diinterpretasikan oleh tenaga kesehatan.

Secara keseluruhan, hasil penelitian menunjukkan bahwa implementasi Machine Learning dalam Biomedical Informatics merupakan pendekatan yang efektif untuk memprediksi risiko hipertensi pada lansia. Model yang dikembangkan memiliki potensi untuk diintegrasikan ke dalam sistem informasi rumah sakit atau aplikasi mobile health, sehingga mendukung deteksi dini, pemantauan berkelanjutan, dan pengambilan keputusan klinis yang lebih cepat, akurat, dan berbasis bukti.

f. Perbandingan Dengan Penelitian Terdahulu

Bandingkan hasil dengan studi terdahulu yang benar-benar relevan, termasuk tinjauan machine learning untuk hipertensi dan studi yang menggunakan EHR. Perbandingan sebaiknya mencakup populasi, ukuran dan sumber dataset, definisi outcome, fitur yang digunakan, desain validasi, algoritma, serta metrik. Perbedaan angka akurasi tidak boleh dianggap sebagai keunggulan langsung jika definisi target, prevalensi kelas, atau metodepembagian data berbeda. Referensi yang telah tercantum dalam naskah—misalnya du Toit et al. (2023), Xu et al. (2021), dan tinjauan eBioMedicine (2023)—dapat ditelaah kembali untuk memastikan kesesuaian isi dan detail bibliografinya sebelum digunakan sebagai dasar perbandingan

IV. KESIMPULAN

Penelitian ini berhasil mengimplementasikan Machine Learning dalam kerangka Biomedical Informatics untuk memprediksi risiko hipertensi pada lansia berdasarkan data Electronic Health Record (EHR). Tahapan penelitian meliputi proses data preprocessing, seleksi fitur, pembangunan model prediksi menggunakan beberapa algoritma Machine Learning, serta evaluasi kinerja model menggunakan metrik Accuracy, Precision, Recall, F1-Score, dan Area Under the Curve (AUC).

Hasil penelitian menunjukkan bahwa algoritma Extreme Gradient Boosting (XGBoost) memberikan performa terbaik dibandingkan algoritma lain yang diuji, dengan nilai Accuracy sebesar 95,16%, Precision sebesar 94,82%, Recall sebesar 94,57%,

e-ISSN : 2964-5115
p-ISSN : 2964-4364

F1-Score sebesar 94,69%, dan AUC sebesar 0,98. Temuan ini menunjukkan bahwa algoritma XGBoost mampu membangun model prediksi yang akurat dalam mengidentifikasi lansia yang berisiko mengalami hipertensi Hasil analisis juga mengidentifikasi bahwa tekanan darah sistolik, tekanan darah diastolik, usia, indeks massa tubuh (BMI), dan riwayat diabetes merupakan faktor yang paling berpengaruh terhadap prediksi risiko hipertensi. Integrasi berbagai variabel klinis tersebut melalui pendekatan Biomedical Informatics memungkinkan pemanfaatan data rekam medis elektronik secara lebih optimal untuk menghasilkan informasi yang bernilai dalam proses pengambilan keputusan klinis.

Dari sisi implementasi, model yang dikembangkan berpotensi diintegrasikan ke dalam Clinical Decision Support System (CDSS) maupun aplikasi Mobile Health (mHealth) untuk mendukung

deteksi dini, skrining risiko, pemantauan kesehatan secara berkelanjutan, serta penyusunan strategi pencegahan hipertensi pada populasi lansia. Penerapan sistem ini diharapkan dapat meningkatkan efektivitas layanan kesehatan, membantu tenaga medis dalam menentukan prioritas intervensi, serta mendukung pengelolaan penyakit kronis yang lebih cepat, tepat, dan berbasis data.

Meskipun demikian, penelitian ini masih memiliki keterbatasan, terutama pada penggunaan dataset yang berasal dari satu institusi layanan kesehatan dan jumlah variabel yang belum sepenuhnya mencakup seluruh faktor risiko hipertensi. Oleh karena itu, penelitian selanjutnya disarankan menggunakan dataset multisenter dengan jumlah data yang lebih besar dan beragam, mengintegrasikan data dari Internet of Medical Things (IoMT) atau perangkat wearable, serta menerapkan pendekatan Explainable Artificial Intelligence (XAI) agar hasil prediksi lebih transparan, mudah dipahami oleh tenaga kesehatan, dan lebih dapat dipercaya dalam praktik klinis.

Secara keseluruhan, penelitian ini membuktikan bahwa penerapan Machine Learning dalam Biomedical Informatics merupakan pendekatan yang efektif untuk mendukung prediksi risiko hipertensi pada lansia. Pendekatan ini tidak hanya meningkatkan akurasi analisis data kesehatan, tetapi juga memberikan kontribusi terhadap pengembangan sistem kesehatan digital yang bersifat prediktif, preventif, personal, dan berbasis bukti (evidence-based healthcare), sehingga dapat menjadi landasan bagi pengembangan layanan kesehatan lansia yang lebih berkualitas di masa mendatang.

V. REFERENSI

du Toit, C., Tran, T. Q. B., Deo, N., Aryal, S., Lip, S., Sykes, R., et al. (2023). *Survey and evaluation of hypertension machine learning research*. Journal of the American Heart Association, 12**(9),** e027896. https://doi.org/10.1161/JAHA.122.027896

Gudigar, A., Kadri, N. A., Raghavendra, U., Samanth, J., Maithri, M., Inamdar, M. A., et al. (2024). *Automatic identification of hypertension and assessment of its secondary effects using artificial intelligence: A systematic review (2013–2023).* Computers in Biology and Medicine, 172**,** 108207. https://doi.org/10.1016/j.compbiomed.2024.108207

Kohjitani, H., Koshimizu, H., Nakamura, K., & Okuno, Y. (2024). *Recent developments in machine learning modeling methods for hypertension treatment*. Journal of Human Hypertension, 47**(3),** 700–707. https://doi.org/10.1038/s41440-023-01547-w

Ye, X., Zeng, Q. T., Facelli, J. C., Brixner, D. I., Conway, M., & Bray, B. E. (2020). *Predicting optimal hypertension treatment pathways using recurrent neural networks*. International Journal of Medical Informatics, 139**,** 104122. https://doi.org/10.1016/j.ijmedinf.2020.104122

Xu, Y., Li, Y., Liu, X., et al. (2021). *An application of machine learning to etiological diagnosis of secondary hypertension: Retrospective study using electronic medical records*. JMIR Medical Informatics, 9**(1),** e19739. https://doi.org/10.2196/19739

eBioMedicine. (2023). *Machine learning in hypertension: From diagnosis to personalised medicine*. eBioMedicine, 92**,** 104658. https://doi.org/10.1016/j.ebiom.2023.104658

e-ISSN : 2964-5115
p-ISSN : 2964-4364

El-Hajj, M., Kyriacou, P. A., & Moussavi, Z. (2021). *A review of machine learning in hypertension detection and blood pressure estimation based on clinical and physiological data*. Biomedical Signal Processing and Control, 68**,** 102813. https://doi.org/10.1016/j.bspc.2021.102813

Hripcsak, G., & Albers, D. J. (2013). *Next-generation phenotyping of electronic health records*. Journal of the American Medical Informatics Association, 20**(1)**, 117–121.

Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). *Mining electronic health records: Towards better research applications and clinical care*. Nature Reviews Genetics, 13(6), 395–405.

Shortliffe, E. H., & Cimino, J. J. (Eds.). (2021). *Biomedical Informatics: Computer Applications in Health Care and Biomedicine* (5th ed.). Springer.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2021). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed., corrected printing). Springer.

Géron, A. (2023). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (3rd ed.). O'Reilly Media.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).

World Health Organization. (2023). *Hypertension*. Geneva: World Health Organization.

World Health Organization. (2022). *Decade of Healthy Ageing: Baseline Report*. Geneva: World Health Organization.

Kementerian Kesehatan Republik Indonesia. (2023). *Profil Kesehatan Indonesia 2023*. Jakarta: Kementerian Kesehatan Republik Indonesia.

Kementerian Kesehatan Republik Indonesia. (2023). *Pedoman Nasional Pelayanan Kedokteran Tata Laksana Hipertensi*. Jakarta: Kementerian Kesehatan Republik Indonesia.