

ANALISIS KOMPARATIF TIGA ALGORITMA MACHINE LEARNING DALAM PREDIKSI EFEKTIVITAS TERAPI ALTERNATIF PASIEN STROKE

Zulfati Dinul Fatiha^{*}, Dicky Octaviano²

^{1,2}Institusi: Universitas Bina Sarana Informatika alamat; Jalan Ciledug Raya No. 108, Cipulir, Jakarta Selatan

Keywords:

Prediksi Pengobatan
Alternatif Stroke, Komparasi
Model Machine Learning,
Random Forest, XGBoost,
Klasifikasi Medis.

Correspondent Email:

zulfati.zdf@bsi.ac.id

Abstrak. Pengobatan alternatif berbasis *Traditional Chinese Medicine* seperti akupunktur, bekam, dan herbal banyak digunakan dalam rehabilitasi stroke, namun efektivitasnya bervariasi antar pasien. Penelitian ini bertujuan membandingkan kinerja tiga algoritma *machine learning* yaitu *Logistic Regression*, *Random Forest*, dan *XGBoost* dalam memprediksi keberhasilan pengobatan alternatif pada pasien stroke. Dataset terdiri dari 1040 rekam medis pasien dengan 15 fitur klinis dari pusat kesehatan di Tangerang. Metode penelitian meliputi preprocessing data, pembagian data 80:20, pelatihan model, dan evaluasi menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan *Cohen's Kappa*. Hasil menunjukkan *Random Forest* mencapai kinerja terbaik dengan akurasi 93,9%, mengungguli *XGBoost* (92,31%) dan *Logistic Regression* (88,1%). *Random Forest* juga menunjukkan *robustness* terhadap konfigurasi default dibandingkan *XGBoost* yang memerlukan hyperparameter tuning intensif untuk performa optimal. Temuan ini merekomendasikan *Random Forest* sebagai pilihan praktis untuk implementasi sistem prediksi klinis dengan sumber daya terbatas. Penelitian berkontribusi pada pengembangan sistem pendukung keputusan berbasis data di layanan pengobatan alternatif stroke.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. Alternative medicine based on *Traditional Chinese Medicine* such as acupuncture, cupping, and herbs is widely used in stroke rehabilitation, but its effectiveness varies among patients. This study aims to compare the performance of three machine learning algorithms—*Logistic Regression*, *Random Forest*, and *XGBoost*—in predicting the success of alternative medicine for stroke patients. The dataset consists of 1040 medical records with 15 clinical features from a health center in Tangerang. Research methods include data preprocessing, 80:20 data splitting, model training, and evaluation using accuracy, precision, recall, *F1-score*, and *Cohen's Kappa* metrics. Results show that *Random Forest* achieved the best performance with 93.91% accuracy, outperforming *XGBoost* (92.31%) and *Logistic Regression* (88.14%). *Random Forest* also demonstrated robustness with default configuration compared to *XGBoost*, which requires intensive hyperparameter tuning for optimal performance. These findings recommend *Random Forest* as a practical choice for clinical prediction system implementation with limited resources. The research contributes to the development of data-based decision support systems in stroke alternative medicine services

1. PENDAHULUAN

Pengobatan alternatif, terutama yang berbasis Traditional Chinese Medicine (TCM) seperti akupunktur, bekam, dan herbal, masih banyak digunakan dalam penanganan berbagai penyakit, termasuk stroke. Pemilihan terapi ini didorong oleh biaya yang relatif terjangkau, efek samping yang minimal, serta ketersediaannya di daerah dengan fasilitas kesehatan terbatas [1], [2]. Pada kasus stroke, beberapa studi melaporkan bahwa akupunktur dapat mendukung pemulihan fungsi motorik dan mengurangi gejala sisa, meskipun respons pasien cenderung bervariasi [2], [1].

Variabilitas hasil terapi tersebut menunjukkan perlunya pendekatan yang lebih objektif dan terukur dalam memprediksi keberhasilan pengobatan alternatif.

Dalam konteks ini, pemanfaatan *machine learning* (ML) menjadi salah satu solusi potensial. ML memungkinkan analisis pola kompleks dari data rekam medis, termasuk faktor-faktor seperti durasi penyakit, jumlah terapi, dan karakteristik klinis pasien [3], [4]. Beberapa penelitian terdahulu telah mengaplikasikan ML untuk memprediksi hasil terapi stroke menggunakan data neuroimaging dan karakteristik pasien [5], [6]. Namun, studi-studi tersebut umumnya masih menggunakan satu algoritma tertentu dan dataset dengan cakupan terbatas.

Sebagian besar penelitian berbasis ML dalam konteks pengobatan alternatif telah berfokus pada prediksi efektivitas terapi akupunktur dengan memanfaatkan data hasil pemeriksaan medis seperti MRI dan EEG [6]–[10]. Misalnya, penelitian untuk memprediksi prognosis pasien depresi yang menjalani akupunktur [8], efektivitas akupunktur pada migrain tanpa aura [6], serta respons terapi pada pasien dismenorea primer [10]. Dataset yang digunakan pada studi-studi tersebut bersumber dari hasil pemeriksaan neuroimaging dan fisiologis, sedangkan dataset berbasis rekam medis klinis dalam format TCM masih jarang dieksplorasi.

Dalam penelitian sebelumnya, salah satu model *machine learning* yaitu *Support Vector Machine* (SVM) merupakan model yang paling sering digunakan untuk memproses dataset pengobatan alternatif, seperti pada prediksi efektivitas akupunktur untuk migrain [6], nyeri punggung kronis [7], dan pemulihan pasca stroke [5]. Meskipun model SVM menunjukkan

kinerja yang baik, model berbasis ensemble seperti model *Random Forest* dan model *XGBoost* telah terbukti unggul dalam berbagai tantangan prediksi medis karena kemampuannya menangani hubungan non-linear dan data tidak seimbang [10], [11].

Pendekatan studi komparatif antaralgoritma ML telah banyak diterapkan dalam penelitian medis untuk mengidentifikasi model dengan kinerja terbaik. Misalnya, penelitian oleh Alfajr dkk. [9] membandingkan *Random Forest* dan *Support Vector Machine* dalam prediksi penyakit jantung, sedangkan Andryan dkk. [10] menguji *XGBoost* dan SVM untuk diagnosis kanker payudara. Hasil dari penelitian tersebut menunjukkan bahwa performa model sangat bergantung pada karakteristik data dan konteks klinis yang dikaji.

Kemudian penelitian sebelumnya yang dilakukan oleh Fatiha dan Subekti [11], telah mengembangkan model prediksi hasil pengobatan alternatif pada pasien stroke menggunakan *machine learning* dengan pendekatan *explainable artificial intelligence* (XAI). Penelitian tersebut menunjukkan potensi *machine learning* dalam membantu memahami faktor-faktor yang memengaruhi keberhasilan terapi.

Dalam domain pengobatan alternatif khususnya stroke, studi komparatif antaralgoritma ML masih sangat terbatas. Padahal, pemilihan model yang tepat dapat meningkatkan akurasi prediksi dan mendukung pengambilan keputusan klinis yang lebih baik. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja tiga algoritma *Machine learning* yaitu *Logistic Regression*, *Random Forest*, dan *XGBoost*, dalam memprediksi hasil pengobatan alternatif pada pasien stroke menggunakan dataset rekam medis.

Dataset yang digunakan terdiri dari 1040 rekam medis pasien stroke yang menjalani terapi akupunktur, bekam, atau herbal di pusat kesehatan di Tangerang, dengan 15 fitur klinis dan demografis. Evaluasi dilakukan berdasarkan metrik akurasi, presisi, *recall*, dan *F1-score* dan skor kappa.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk melakukan studi komparatif terhadap beberapa algoritma *machine learning*, yaitu *Logistic Regression*, *Random Forest*, dan *XGBoost*, dalam

memprediksi hasil pengobatan alternatif menggunakan data rekam medis pasien.

Evaluasi dilakukan menggunakan metrik performa untuk menentukan model dengan kinerja terbaik. Hasil penelitian ini diharapkan dapat memberikan gambaran objektif mengenai model yang paling optimal dalam memprediksi keberhasilan pengobatan alternatif serta menjadi dasar pengembangan sistem pendukung keputusan berbasis data di bidang kesehatan alternatif.

2. TINJAUAN PUSTAKA

2.1 Pengobatan Alternatif pada Stroke

Pengobatan alternatif berbasis *Traditional Chinese Medicine* (TCM) seperti akupunktur, bekam, dan herbal telah lama digunakan sebagai pendukung rehabilitasi stroke. Penelitian terapi Akupunktur pada keluhan Stroke, dilaporkan dapat meningkatkan aliran darah serebral, mengurangi spasme otot, dan memodulasi aktivitas neural yang mendukung pemulihan fungsi motorik [2], [1].

Studi oleh Lu et al. [5] menunjukkan bahwa akupunktur dapat menginduksi reorganisasi fungsional otak pasca stroke yang dapat diprediksi menggunakan pendekatan *machine learning* berbasis fMRI. Meski demikian, variasi respons klinis antarpasien menunjukkan bahwa faktor seperti durasi penyakit, frekuensi terapi, dan kondisi komorbid turut memengaruhi hasil akhir terapi [1], [2].

2.2 Machine Learning dalam Prediksi Kesehatan

Pemanfaatan machine learning dalam bidang kesehatan semakin meluas, terutama untuk analisis data rekam medis elektronik (*electronic health records* / EHR) dan prediksi hasil terapi [3], [4]. Algoritma ML mampu mengidentifikasi pola kompleks dan hubungan non-linear antar variabel klinis yang sulit dikenali dengan metode statistik konvensional [4]. Dalam konteks stroke, ML telah digunakan untuk memprediksi pemulihan motorik, risiko rekurensi, dan respons terhadap intervensi terapeutik [5], [6].

2.3 Algoritma yang Dibandingkan

2.3.1 Logistic Regression

Logistic Regression merupakan model linier yang umum digunakan untuk klasifikasi biner dalam penelitian medis karena

kesederhanaan dan kemudahan interpretasi [12]. Model ini memperkirakan probabilitas hasil berdasarkan fungsi logit, namun memiliki keterbatasan dalam menangani hubungan *non-linear* dan interaksi kompleks antar fitur [12].

2.3.2 Random Forest

Random Forest adalah algoritma ensemble berbasis pohon keputusan yang bekerja dengan membangun banyak pohon dan mengagregasi hasilnya. Metode ini dikenal *robust* terhadap *overfitting* dan mampu menangani data tidak seimbang serta fitur dengan skala berbeda [13]. Dalam berbagai studi kesehatan, *Random Forest* sering menunjukkan kinerja yang stabil dan akurat [10], [13].

2.3.3 XGBoost

XGBoost merupakan pengembangan dari algoritma *Gradient Boosting* yang dioptimalkan untuk kecepatan dan kinerja. Algoritma ini membangun model secara sekuensial dengan memperbaiki kesalahan model sebelumnya, sehingga sering mencapai akurasi tinggi dalam tantangan prediksi medis [10], [14]. *XGBoost* juga memiliki mekanisme regulasi untuk mencegah *overfitting* dan mendukung penanganan data *missing value* [14].

2.4. Studi Komparatif Algoritma dalam Kesehatan

Studi komparatif antaralgoritma ML penting untuk menentukan model yang paling sesuai dengan karakteristik data tertentu. Alfajr et al. [9] membandingkan *Random Forest* dan SVM dalam prediksi penyakit jantung, menyimpulkan bahwa *Random Forest* lebih unggul dalam akurasi dan stabilitas. Sementara itu, Andryan et al. [10] menemukan bahwa *XGBoost* mengungguli SVM dalam diagnosis kanker payudara. Namun, penelitian yang membandingkan *Logistic Regression*, *Random Forest*, dan *XGBoost* secara bersamaan dalam konteks prediksi hasil pengobatan alternatif stroke masih sangat terbatas.

2.5 Penelitian Terkait

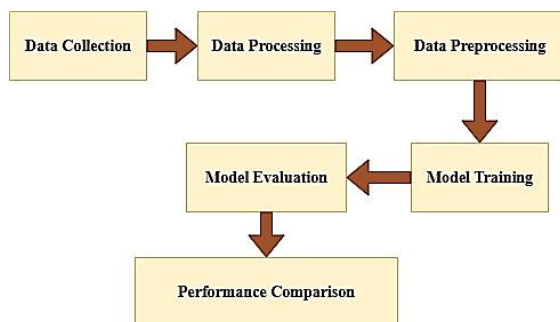
Penelitian yang terkait sebelumnya menerapkan pendekatan *machine learning* untuk memprediksi hasil beberapa pengobatan alternatif untuk menentukan prognosis pada pasien stroke dengan menambahkan aspek interpretabilitas model [11]. Namun penelitian

tersebut belum membandingkan kinerja antar algoritma secara khusus.

Selain itu, banyak penelitian sebelumnya mengenai pengobatan alternatif cenderung hanya menggunakan satu algoritma atau dataset yang terbatas [6], [8]. Oleh karena itu, penelitian ini menerapkan studi komparatif terhadap beberapa algoritma *machine learning* guna menemukan model yang paling efektif dalam memprediksi hasil pengobatan alternatif.

3. METODE PENELITIAN

Penelitian ini menerapkan pendekatan *comparative experimental study* dengan tujuan mengevaluasi dan membandingkan kinerja beberapa algoritma *machine learning* dalam memprediksi prognosis pengobatan alternatif terhadap pasien stroke.



Gambar 1 Metode Penelitian

Tahapan penelitian meliputi *data collection*, *data processing*, *data preprocessing*, *model training*, *model evaluation*, dan *performance comparison* atau perbandingan kinerja model (Gambar 1). Perbandingan tersebut dilakukan berdasarkan metrik evaluasi kinerja klasifikasi untuk menentukan model yang paling optimal.

3.1 Data Collection

Data yang digunakan dalam penelitian ini diperoleh dari rekam medis pasien pengobatan alternatif di salah satu pusat layanan kesehatan di wilayah Tangerang, Indonesia. Data dikumpulkan selama periode sepuluh tahun, sehingga mencerminkan variasi karakteristik pasien dan terapi yang cukup beragam.

Dataset mencakup pasien yang menjalani beberapa jenis pengobatan alternatif, antara lain akupunktur, bekam, dan penggunaan herbal. Dataset tersebut terdiri dari 1.040 data (*instance*) dengan 15 atribut (*fitur*), yang

meliputi usia, pekerjaan, jenis penyakit, status hipertensi, lama menderita penyakit, jumlah terapi, durasi terapi, jenis pengobatan alternatif, serta prognosis pasien sebagai label keluaran (*output*). Seluruh data telah melalui proses anonimisasi untuk menjaga kerahasiaan identitas pasien dan digunakan secara eksklusif untuk kepentingan penelitian.

3.2 Data Processing

Dataset yang digunakan terlebih dahulu diidentifikasi dan diseleksi untuk memastikan kesesuaiannya dengan tugas prediksi hasil pengobatan alternatif. Pada tahap ini dilakukan proses *data processing* yang meliputi pembersihan data dan seleksi fitur agar dataset dapat diproses secara optimal oleh model klasifikasi.

Beberapa atribut penting terpaksa tidak digunakan karena memiliki proporsi nilai hilang yang sangat tinggi, yaitu sekitar 70–80%, seperti kondisi emosional pasien, kondisi fisik pasien secara subjektif, serta kondisi lidah pasien. Penghapusan atribut tersebut dilakukan untuk menjaga kualitas data dan menghindari bias pada proses pelatihan model [15].

3.3 Data Preprocessing

Tahap pra-pemrosesan data dilakukan untuk meningkatkan kualitas data sebelum digunakan dalam proses pelatihan model. Tahapan *preprocessing* meliputi:

- Pembersihan data, yaitu penghapusan data duplikat dan penanganan nilai yang hilang pada fitur yang masih digunakan dengan metode *mean imputation* untuk data numerik dan *mode imputation* untuk data kategorikal.
- Transformasi data kategorikal ke dalam bentuk numerik menggunakan teknik *label encoding* dan *one-hot encoding* agar dapat diproses oleh algoritma *machine learning* [15].
- Normalisasi data numerik dengan *min-max scale* untuk menyamakan skala antar fitur. Metode ini digunakan untuk memastikan seluruh fitur numerik berada pada rentang yang sama sehingga proses pelatihan model menjadi lebih stabil [15]. Normalisasi data diterapkan pada *Logistic Regression* karena sensitif terhadap skala fitur. Pada *Random Forest* dan *XGBoost*, normalisasi tidak wajib, namun tetap dilakukan untuk menjaga konsistensi proses pemodelan.

e. Pembagian dataset menjadi data latih dan data uji dengan rasio 80:20, yang merupakan proporsi umum dalam penelitian *machine learning* untuk menjaga keseimbangan antara pelatihan dan evaluasi model [16].

3.4 Model Training

Pada tahap pelatihan model, digunakan tiga algoritma *machine learning* yang mewakili pendekatan berbeda, yaitu:

a. Logistic Regression (LR)

Logistic Regression digunakan sebagai model baseline karena memiliki struktur yang sederhana dan interpretasi yang jelas, serta sering digunakan sebagai pembanding awal dalam penelitian klasifikasi medis [12].

b. Random Forest (RF)

Random Forest merupakan metode ensemble berbasis decision tree yang mampu menangani hubungan non-linear dan mengurangi risiko overfitting melalui mekanisme bootstrap aggregation [13].

c. XGBoost (Extreme Gradient Boosting)

XGBoost adalah algoritma boosting yang dikenal memiliki performa tinggi dalam berbagai tugas klasifikasi dan prediksi di bidang kesehatan, terutama pada dataset dengan pola kompleks [10].

Ketiga model dilatih menggunakan data latih yang sama dengan implementasi menggunakan pustaka Scikit-learn (untuk *Logistic Regression* dan *Random Forest*) dan *XGBoost* library (untuk *XGBoost*) untuk menjamin evaluasi performa yang seimbang antar algoritma.

3.5 Model Evaluation

Evaluasi kinerja perbandingan model dilakukan menggunakan data uji yang tidak terlibat dalam proses pelatihan. Beberapa metrik evaluasi digunakan untuk menilai performa model secara komprehensif, yaitu:

- Accuracy*, untuk mengukur tingkat ketepatan prediksi secara keseluruhan.
- Precision*, *Recall*, dan *f1-score*, untuk mengevaluasi keseimbangan antara prediksi positif dan negatif.
- Cohen's Kappa*, untuk mengukur tingkat kesepakatan antara hasil prediksi model dan data aktual dengan mempertimbangkan kemungkinan kesepakatan secara acak [18].

Penggunaan berbagai metrik evaluasi penting dalam konteks data medis karena satu metrik saja tidak selalu cukup untuk menggambarkan performa model secara menyeluruh. Selain itu, analisis matriks konfusi digunakan untuk memahami pola kesalahan klasifikasi pada setiap model.

3.6 Performance Comparison

Tahap akhir dalam metode penelitian ini adalah perbandingan performa ketiga model berdasarkan metrik evaluasi yang telah ditentukan. Hasil perbandingan digunakan untuk mengidentifikasi model yang paling optimal dalam memprediksi keberhasilan pengobatan alternatif. Analisis ini menjadi dasar untuk pembahasan lebih lanjut mengenai kelebihan dan keterbatasan masing-masing algoritma pada bagian hasil dan pembahasan.

3.7 Tools dan Environment

Penelitian ini dilaksanakan menggunakan Python 3.9 dengan pustaka utama Scikit-learn 1.3, *XGBoost* 1.7, Pandas, dan NumPy, serta dijalankan pada platform *Google Colaboratory*.

4. HASIL DAN PEMBAHASAN

4.1 Hasil Eksperimen

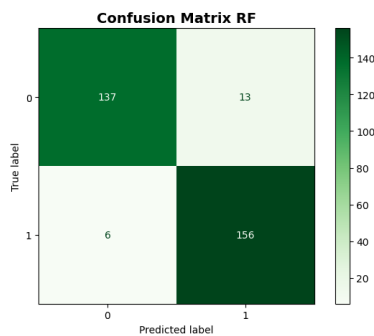
Berikut hasil eksperimen disajikan dalam Tabel 1 yang menampilkan perbandingan kinerja ketiga model *machine learning* dalam memprediksi keberhasilan pengobatan alternatif pada pasien stroke. Evaluasi dilakukan menggunakan berbagai metrik untuk mendapatkan gambaran komprehensif tentang performa masing-masing algoritma.

Tabel 1. Perbandingan Kinerja Model Prediksi Efektivitas Pengobatan Alternatif

| Model | Akurasi | Presisi | Recall | F1-Score |
|-------|---------|---------|--------|----------|
| LR | 0,88 | 0,88 | 0,88 | 0,88 |
| RF | 0,93 | 0,93 | 0,93 | 0,93 |
| XGB | 0,92 | 0,92 | 0,92 | 0,92 |

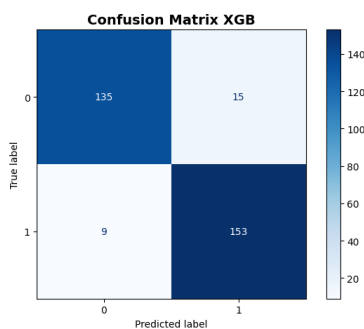
Berdasarkan Tabel 1, *Random Forest* (RF) menunjukkan kinerja terbaik dengan akurasi tertinggi sebesar 93,91%, diikuti oleh *XGBoost* (XGB) dengan akurasi 92,31%, dan *Logistic Regression* (LR) dengan akurasi 88,14%. Pola yang sama terlihat pada semua metrik evaluasi,

di mana *Random Forest* konsisten mengungguli kedua model lainnya.



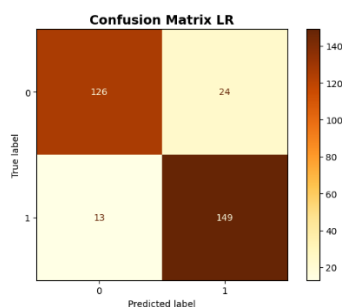
Gambar 2. *Confusion Matrix Random Forest*

Analisis lebih mendalam melalui matriks konfusi pada model *Random Forest* yang ditampilkan pada Gambar 2 mengungkapkan pola kesalahan klasifikasi masing-masing model. *Random Forest* menghasilkan 6 false negative dan 13 false positive dari total 312 sampel uji.



Gambar 3. *Confusion Matrix XGBoost*

Sementara itu, *Confusion Matrix* pada model *XGBoost* (Gambar 3) menghasilkan 9 false negative dan 15 false positive, dan *Logistic Regression* menghasilkan 13 false negative dan 24 false positive.



Gambar 4. *Confusion Matrix Logistic Regression*

Nilai *Cohen's Kappa* untuk *Random Forest* sebesar 0,87 menunjukkan tingkat kesepakatan yang hampir sempurna antara prediksi model dengan label aktual, sementara nilai untuk *Logistic Regression* (0,76) mengindikasikan kesepakatan yang substansial (Tabel 2).

Tabel 2. Perbandingan *Cohen's Kappa* pada model komparasi untuk prognosis pasien stroke

| Model | Cohen's Kappa |
|---------------------|---------------|
| Logistic Regression | 0.7618 |
| Random Forest | 0.8778 |
| XGBoost | 0.8457 |

4.2 Pembahasan

4.2.1 Interpretasi Hasil dan Keunggulan *Random Forest*

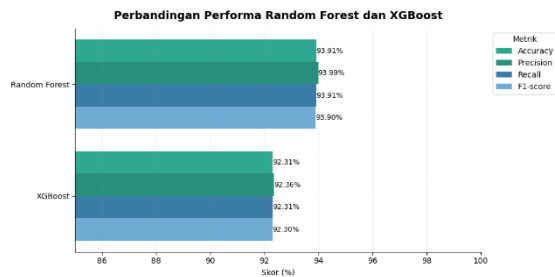
Berdasarkan analisis yang dilakukan, *Random Forest* terbukti menjadi model paling akurat untuk memprediksi hasil terapi alternatif pada stroke. Keefektifan model ini didukung oleh beberapa prinsip algoritma. Faktor pertama, kemampuannya menangkap interaksi non-linier yang kompleks di antara fitur-fitur klinis, seperti panjangnya riwayat penyakit, intensitas terapi, dan metode intervensi berkat konstruksinya sebagai model ensemble yang terdiri dari banyak pohon keputusan [17]. Kedua, mekanisme bootstrap aggregation (bagging) pada *Random Forest* efektif mengurangi varians dan mencegah overfitting, yang sangat penting dalam konteks data medis yang seringkali memiliki noise tinggi [13].

4.2.2 Analisis Perbandingan *Random Forest* vs *XGBoost*

Analisis komparatif mendalam terhadap kinerja *Random Forest* (RF) dan *XGBoost* (XGB) dalam penelitian ini mengungkap temuan yang signifikan dan patut diperhatikan. Meskipun *XGBoost* secara luas diakui dalam literatur sebagai salah satu algoritma *boosting* paling canggih, yang kerap mencatat performa tinggi dalam beragam tantangan prediksi di bidang kedokteran dan kesehatan [14], hasil eksperimen yang dilakukan justru menunjukkan fenomena sebaliknya.

Dalam konteks dataset pengobatan alternatif stroke yang digunakan, model *Random Forest* berhasil meraih akurasi sebesar 93,91%, mengungguli akurasi *XGBoost* yang mencapai 92,31%. Keunggulan RF tidak hanya

tampak pada metrik akurasi, namun juga konsisten terlihat pada metrik evaluasi lain seperti presisi, *recall*, dan *F1-score*, yang kesemuanya menunjukkan nilai lebih tinggi dibandingkan XGB. Visualisasi perbandingan menyeluruh antara kedua model berdasarkan keempat metrik kunci tersebut dapat diamati secara detail pada Gambar 5.



Gambar 5. Perbandingan performa *Random Forest* dan *XGBoost* berdasarkan metrik akurasi, presisi, *recall*, dan *F1-score*.

Perbedaan ini dapat dijelaskan melalui beberapa faktor yaitu ketahanan konfigurasi *Default* yang dimana pada model *Random Forest* menunjukkan *robustness* yang lebih baik dengan konfigurasi default, sementara model *XGBoost* sangat sensitif terhadap pengaturan hyperparameter [14]. Dalam penelitian ini, kedua model diuji dengan konfigurasi standar tanpa hyperparameter tuning intensif.

Hasil ini berbeda dengan penelitian sebelumnya oleh Fatiha dan Subekti [11] yang melaporkan *XGBoost* mencapai akurasi 94,7%. Perbedaan signifikan ini disebabkan oleh variasi dalam proses optimasi *hyperparameter*. Penelitian sebelumnya melakukan tuning intensif pada *XGBoost*, sementara dalam penelitian ini digunakan konfigurasi standar untuk ketiga model.

Karakteristik utama *Random Forest* terletak pada mekanisme *bagging*-nya, di mana berbagai pohon keputusan dibangun secara terpisah dan hasilnya digabungkan. Ini menjadikan model tersebut lebih stabil dalam menangani data yang mengandung *noise* [17]. Sedangkan *XGBoost* menggunakan pendekatan *boosting* berurutan yang pada dasarnya lebih peka terhadap *noise* dan menuntut penyetelan parameter yang lebih presisi.

4.2.3 Performa *Logistic Regression* sebagai Model *Baseline*

Logistic Regression menunjukkan performa yang wajar dengan akurasi 88,14%, namun secara signifikan lebih rendah dibandingkan kedua model *ensemble*. Hasil ini mengkonfirmasi bahwa hubungan antara fitur-fitur klinis dengan hasil pengobatan alternatif bersifat non-linear dan kompleks, sehingga tidak dapat dimodelkan secara optimal dengan pendekatan linier seperti *Logistic Regression* [12]. Meski demikian, *Logistic Regression* tetap memiliki nilai sebagai model *baseline* yang mudah diinterpretasi dan dapat memberikan wawasan awal tentang hubungan antar variabel.

4.2.4 Implikasi Terhadap Pemilihan Model dalam Konteks Klinis

Temuan penelitian ini memiliki implikasi penting untuk pemilihan algoritma dalam implementasi klinis.

Untuk Setting dengan Sumber Daya Terbatas: *Random Forest* dengan konfigurasi default menjadi pilihan yang lebih praktis dan efisien, karena mencapai performa tinggi tanpa memerlukan tuning intensif.

Untuk Lingkungan dengan Kapasitas Optimasi: *XGBoost* tetap memiliki potensi performa lebih tinggi setelah melalui proses *hyperparameter* tuning yang komprehensif, seperti yang ditunjukkan dalam penelitian terdahulu [11].

Pertimbangan Stabilitas versus Performa Maksimal: Pilihan antara *Random Forest* dan *XGBoost* harus mempertimbangkan *trade-off* antara stabilitas prediksi (*Random Forest*) dan potensi performa maksimal (*XGBoost* dengan *tuning*).

4.2.5 Perbandingan dengan Penelitian Terdahulu di Domain Serupa

Temuan penelitian ini selaras dengan beberapa studi komparatif algoritma di domain kesehatan. Penelitian Alfajr dkk. [9] juga menemukan *Random Forest* mengungguli SVM dalam prediksi penyakit jantung.

Begitu pula dengan penelitian Andryan dkk. [10] yang melaporkan keunggulan *XGBoost* terhadap SVM dalam diagnosis kanker payudara. Variasi hasil ini memperkuat kesimpulan bahwa kinerja algoritma sangat kontekstual dan bergantung pada karakteristik spesifik dataset serta domain masalah [11].

4.2.6 Implikasi Klinis dan Implementasi

Tingginya akurasi *Random Forest* (93,91%) menunjukkan potensi aplikasi klinis yang signifikan. Model ini dapat diintegrasikan ke dalam sistem pendukung keputusan klinis di fasilitas layanan pengobatan alternatif untuk:

- Personalisasi terapi berdasarkan profil pasien individu.
- Optimasi alokasi sumber daya dengan memprediksi pasien yang paling mungkin mendapat manfaat dari terapi tertentu.
- Konseling pasien dan keluarga yang lebih terinformasi tentang prognosis pengobatan

Namun, implementasi klinis memerlukan pertimbangan lebih lanjut mengenai interpretabilitas model. Meskipun *Random Forest* kurang mudah diinterpretasi dibandingkan *Logistic Regression*, teknik explainable AI seperti *feature importance* dari *Random Forest* dapat memberikan wawasan tentang kontribusi relatif masing-masing fitur terhadap prediksi.

4.2.7 Keterbatasan dan Saran Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diakui:

- Dataset terbatas pada satu pusat layanan, sehingga generalisasi temuan perlu validasi lebih lanjut.
- Beberapa fitur klinis penting seperti kondisi emosional dan fisik pasien tidak dapat dimanfaatkan karena missing value tinggi.
- Optimasi parameter belum dilakukan secara ekstensif untuk semua model

Untuk penelitian lanjutan, disarankan:

- Validasi eksternal menggunakan dataset dari berbagai pusat layanan Kesehatan.
- Studi komparatif dengan *hyperparameter tuning* pada semua model untuk evaluasi yang lebih adil.
- Eksplorasi teknik *feature imputation* untuk memanfaatkan fitur dengan missing value
- Optimasi *hyperparameter* komprehensif termasuk teknik seperti *grid search* atau *bayesian optimization*

- Integrasi teknik *explainable AI* untuk meningkatkan interpretabilitas prediksi klinis.

4.3 Kesimpulan Bagian Hasil dan Pembahasan

Berdasarkan eksperimen dan analisis yang dilakukan, dapat disimpulkan bahwa:

- Random Forest* menunjukkan kinerja terbaik dengan akurasi 93,91% dalam memprediksi efektivitas pengobatan alternatif pada pasien stroke.
- Random Forest* lebih robust dengan konfigurasi default dibandingkan *XGBoost* dalam konteks dataset ini, menjadikannya pilihan praktis untuk implementasi dengan sumber daya terbatas.
- Performa *XGBoost* sangat bergantung pada optimasi *hyperparameter*, dengan potensi mencapai performa lebih tinggi setelah *tuning* intensif.
- Hubungan antar fitur klinis bersifat non-linear kompleks, sehingga model ensemble seperti *Random Forest* dan *XGBoost* lebih unggul dibandingkan model linier seperti *Logistic Regression*.
- Kinerja algoritma sangat kontekstual dan bergantung pada karakteristik spesifik dataset, sebagaimana dibuktikan dengan perbedaan hasil dibandingkan penelitian terdahulu
- Terdapat potensi implementasi klinis yang signifikan, namun memerlukan validasi lebih lanjut dan pertimbangan aspek interpretabilitas model

Temuan penelitian ini memberikan dasar empiris untuk pemilihan algoritma *machine learning* yang optimal dalam konteks prediksi hasil pengobatan alternatif, khususnya untuk rehabilitasi stroke, serta membuka peluang untuk pengembangan sistem pendukung keputusan klinis berbasis data di masa depan.

5. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan mengenai perbandingan kinerja tiga algoritma machine learning untuk prediksi efektivitas pengobatan alternatif pada pasien stroke, dapat disimpulkan sebagai berikut:

a. Hasil Utama:

Random Forest menunjukkan kinerja terbaik dengan akurasi 93,91%, diikuti oleh *XGBoost* (92,31%) dan *Logistic Regression* (88,14%).

Random Forest konsisten unggul dalam semua metrik evaluasi (presisi, recall, F1-score, dan *Cohen's Kappa*).

b. Kelebihan Model:

Random Forest: Robust dengan konfigurasi default, stabil pada data dengan noise, dan efektif menangkap hubungan non-linear kompleks antar fitur klinis.

XGBoost: Memiliki potensi performa lebih tinggi setelah hyperparameter tuning intensif, seperti yang ditunjukkan penelitian sebelumnya (94,7%).

Logistic Regression: Mudah diinterpretasi dan berguna sebagai baseline model untuk analisis awal.

c. Keterbatasan Penelitian:

Dataset terbatas pada satu pusat layanan sehingga generalisasi temuan memerlukan validasi eksternal. Beberapa fitur klinis penting tidak dapat dimanfaatkan karena missing value tinggi (>70%). *Hyperparameter* tuning belum dilakukan secara ekstensif pada semua model.

d. Implikasi Praktis:

Untuk implementasi klinis dengan sumber daya terbatas, *Random Forest* dengan konfigurasi default merupakan pilihan paling praktis dan efisien. Dalam lingkungan dengan kapasitas komputasi memadai, *XGBoost* dengan tuning intensif memiliki potensi mencapai performa optimal.

e. Rekomendasi Pengembangan Selanjutnya:

- Validasi eksternal menggunakan dataset dari berbagai pusat layanan kesehatan untuk meningkatkan generalisasi model.
- Studi komparatif dengan *hyperparameter* tuning pada semua model untuk evaluasi yang lebih komprehensif dan adil.
- Eksplorasi teknik feature imputation untuk memanfaatkan fitur dengan missing value tinggi.
- Integrasi teknik *explainable AI* (XAI) seperti SHAP atau LIME untuk meningkatkan interpretabilitas prediksi klinis.

Pengembangan sistem pendukung keputusan klinis berbasis model terbaik untuk personalisasi terapi pengobatan alternatif.

Penelitian ini memberikan kontribusi empiris dalam pemilihan algoritma machine learning optimal untuk prediksi efektivitas pengobatan alternatif, khususnya dalam konteks rehabilitasi stroke, serta membuka jalan bagi pengembangan solusi berbasis data yang dapat mendukung pengambilan keputusan klinis yang lebih tepat dan personal.

DAFTAR PUSTAKA

- [1] D. A. Andira and J. K. Pudjibudojo, "Pengobatan Alternatif Sebagai Upaya Penyembuhan Penyakit," *Insight J. Pemikir. dan Penelit. Psikol.*, vol. 16, no. 2, pp. 393–401, 2020.
- [1] D. A. Andira and J. K. Pudjibudojo, "Pengobatan alternatif sebagai upaya penyembuhan penyakit," *Insight J. Pemikir. dan Penelit. Psikol.*, vol. 16, no. 2, pp. 393–401, 2020.
- [2] G. Noor Alivian and K. N. Pratama, "Efektifitas terapi akupunktur terhadap keberhasilan rehabilitasi pasien pasca stroke: Literature review," *J. Bionursing*, vol. 4, no. 1, pp. 29–35, 2022.
- [3] S. Shickel et al., "Deep EHR: A survey of recent advances in deep learning techniques for electronic health record analysis," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 5, pp. 1589–1604, May 2018.
- [4] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New Engl. J. Med.*, vol. 380, no. 14, pp. 1347–1358, Apr. 2019.
- [5] M. Lu et al., "Neuroimaging mechanisms of acupuncture on functional reorganization for post-stroke motor improvement: A machine learning-based fMRI study," *Front. Neurosci.*, vol. 17, pp. 1–18, May 2023.
- [6] T. Yin et al., "The spontaneous activity pattern of the middle occipital gyrus predicts the clinical efficacy of acupuncture treatment for migraine without aura," *Front. Neurol.*, vol. 11, pp. 1–12, Nov. 2020.
- [7] Y. Tu et al., "Multivariate resting-state functional connectivity predicts responses to real and sham acupuncture treatment in chronic low back pain," *NeuroImage Clin.*, vol. 23, p. 101885, Apr. 2019.
- [8] S. Yu et al., "Resting-state functional connectivity patterns predict acupuncture treatment response in primary dysmenorrhea," *Front. Neurosci.*, vol. 14, pp. 1–11, Sep. 2020.

- [9] N. H. Alfajr, G. Garno, and D. Yusup, "Studi komparasi algoritma Random Forest Classifier dan Support Vector Machine dalam prediksi penyakit jantung," *J. Inform. dan Tek. Elektro Terapan*, vol. 13, no. 3, pp. 1–10, 2025, doi: 10.23960/jitet.v13i3.6569.
- [10] M. R. Andryan, M. Fajri, and N. Sulistyowati, "Komparasi kinerja algoritma XGBoost dan Support Vector Machine untuk diagnosis kanker payudara," *JIKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 1, pp. 1–10, Jan. 2022.
- [11] Z. D. Fatiha and A. Subekti, "Explainable prediction of alternative medicine outcome using machine learning and Shapley values," in *Proc. 2023 Int. Conf. Inf. Technol. Res. Innov. (ICITRI)*, Aug. 2023, pp. 1–6, doi: 10.1109/ICITRI59340.2023.10249346.
- [12] H. B. Patel et al., "Logistic regression in clinical and health services research: A comparative study," *J. Med. Syst.*, vol. 45, no. 3, p. 32, Mar. 2021.
- [13] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
- [14] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2016, pp. 785–794.
- [15] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*. Springer, 2015.
- [16] R. A. Sarno and D. A. I. Sensuse, "Evaluasi metode K-fold cross validation untuk klasifikasi data medis menggunakan algoritma C4.5 dan Naïve Bayes," *J. RESTI (Rekayasa Sist. dan Teknol. Inform.)*, vol. 5, no. 2, pp. 284–291, Apr. 2021.
- [17] A. F. Huda and R. Wijaya, "Analisis performa Random Forest pada klasifikasi data tidak seimbang di bidang kesehatan," *J. Ilmu Komput. dan Inform.*, vol. 14, no. 2, pp. 45–56, Aug. 2022.
- [18] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, Mar. 1977.