Journal of Smart Computing and Digital Transformation
Vol. 1, No. 1, Agustus 2026, pp. xx~xx
E-ISSN: 0000-000
DOI: 00.0000/vertex.v2i2.xxxx

Analisis Komparatif Artificial Neural Network Optimasi Particle Swarm Optimization dan Random Forest dengan Recursive Feature Elimination untuk Klasifikasi Diabetes

Muhammad Hamdi[1], Hairani[2], Safrin[3]
Universitas Bumigora, mataram, Indonesia

Article Info

Article history:

Received mm dd, yyyy
Revised mm dd, yyyy
Accepted mm dd, yyyy

Keywords:

Diabetes,
Artificial Neural Network,
Particle Swarm Optimization,
Random Forest,
Recursive Feature Elimination,
Klasifikasi.

ABSTRACT

Diabetes Mellitus merupakan penyakit kronis dengan prevalensi global yang terus meningkat, sehingga memerlukan sistem diagnosis yang akurat dan efisien. Penelitian ini bertujuan melakukan analisis komparatif antara *Artificial Neural Network* (ANN) yang dioptimasi menggunakan *Particle Swarm Optimization* (PSO) dengan *Random Forest* (RF) berbasis *Recursive Feature Elimination* (RFE) untuk klasifikasi diabetes menggunakan Pima Indians Diabetes Dataset. Tahapan penelitian meliputi pemrosesan data, penanganan nilai tidak valid, normalisasi, serta eksperimen pada empat model: ANN *baseline*, ANN-PSO, RF *baseline*, dan RF-RFE. Hasil penelitian menunjukkan bahwa ANN *baseline* memberikan performa terbaik dengan akurasi 81%, *recall* 89%, dan F1-score 82,41%. Optimasi PSO pada ANN belum mampu meningkatkan kinerja pada data pengujian, sementara penerapan RFE pada RF tidak memberikan perubahan performa yang signifikan dibandingkan model RF *baseline*. Temuan ini mengindikasikan bahwa efektivitas teknik optimasi dan seleksi fitur sangat bergantung pada karakteristik data. Penelitian ini berkontribusi dalam pengembangan model klasifikasi medis yang akurat dan dapat menjadi acuan dalam pengambilan keputusan klinis.

Corresponding Author:

Muhammad Hamdi,
Universitas Bumigora, Mataram, Indonesia.
Email: bagloncity@gmail.com.

Cara Mengutip: nama, tahun, "Analisis Sentimen dan Pemodelan Topik Aplikasi Kitabisa menggunakan Support Vector Machine (SVM) dan Metode Tautan Smote-Tomek", **Journal of Smart Computing and Digital Transformation**, vol. 2, no. 2, hlm. 81 - 91, Sep. 2023. doi : 10.30812/ vertex.v 0i0.000 .

1 INTRODUCTION

Diabetes Mellitus merupakan salah satu penyakit kronis dengan tingkat prevalensi yang terus meningkat secara global dan menjadi tantangan serius dalam bidang kesehatan masyarakat [1]. Penyakit ini ditandai oleh tingginya kadar glukosa dalam

darah akibat gangguan produksi insulin maupun resistensi insulin dalam tubuh. Menurut International Diabetes Federation, jumlah penderita diabetes dunia diperkirakan terus mengalami peningkatan signifikan dalam beberapa dekade mendatang seiring perubahan pola hidup, pola makan, serta rendahnya aktivitas fisik masyarakat. Kondisi tersebut menyebabkan diabetes menjadi salah satu penyebab utama komplikasi serius seperti gagal ginjal, penyakit jantung, stroke, dan gangguan penglihatan yang dapat menurunkan kualitas hidup penderita serta meningkatkan angka kematian [2]. Dalam praktik klinis, proses diagnosis diabetes sering kali memerlukan analisis berbagai parameter medis seperti kadar glukosa, indeks massa tubuh, tekanan darah, usia, riwayat keluarga, dan variabel laboratorium lainnya. Kompleksitas hubungan antarvariabel tersebut menyebabkan pendekatan konvensional memiliki keterbatasan dalam menghasilkan diagnosis yang cepat, akurat, dan konsisten [3].

Oleh karena itu, perkembangan teknologi kecerdasan buatan khususnya machine learning dan deep learning mulai banyak dimanfaatkan untuk membantu proses klasifikasi dan prediksi penyakit diabetes secara otomatis. Salah satu metode yang banyak digunakan dalam klasifikasi data medis adalah Artificial Neural Network atau Jaringan Saraf Tiruan (JST). JST memiliki kemampuan yang baik dalam mengenali pola nonlinier dan hubungan kompleks pada data kesehatan sehingga sering menghasilkan performa klasifikasi yang tinggi. Namun demikian, JST juga memiliki beberapa kelemahan, terutama pada proses penentuan bobot awal dan parameter jaringan yang sering menyebabkan model terjebak pada local minima serta menghasilkan proses pelatihan yang kurang optimal [4]. Permasalahan tersebut mendorong pemanfaatan metode optimasi berbasis metaheuristik untuk meningkatkan performa JST. Salah satu algoritma optimasi yang cukup populer adalah Particle Swarm Optimization. PSO bekerja dengan meniru perilaku kawanan partikel dalam mencari solusi optimal pada ruang pencarian tertentu. Dalam implementasinya pada JST, PSO mampu mengoptimalkan bobot dan parameter jaringan sehingga meningkatkan akurasi klasifikasi, mempercepat konvergensi, serta mengurangi risiko [5]. Beberapa penelitian menunjukkan bahwa kombinasi ANN dan PSO mampu menghasilkan performa yang lebih baik dibandingkan ANN konvensional, khususnya pada data tabular medis yang memiliki pola kompleks dan dimensi fitur yang cukup tinggi [6].

Selain pendekatan berbasis neural network, algoritma Random Forest juga menjadi salah satu metode yang banyak digunakan dalam klasifikasi diabetes karena memiliki kemampuan generalisasi yang baik, tahan terhadap noise, serta efektif dalam menangani data multidimensi. Random Forest mampu melakukan proses klasifikasi secara stabil melalui kombinasi banyak decision tree sehingga sering menghasilkan performa yang lebih baik dibandingkan metode klasifikasi konvensional pada data medis [7]. Meskipun demikian, tingginya jumlah fitur pada dataset medis dapat memengaruhi efisiensi komputasi dan performa model sehingga diperlukan proses seleksi fitur yang optimal. Salah satu teknik seleksi fitur yang efektif adalah Recursive Feature Elimination (RFE), yaitu metode seleksi fitur yang bekerja dengan menghapus fitur dengan kontribusi terendah secara bertahap hingga diperoleh subset fitur terbaik. Pendekatan berbasis Random Forest dan RFE dinilai mampu meningkatkan akurasi klasifikasi, mengurangi redundansi data, serta meningkatkan efisiensi komputasi pada data berdimensi tinggi [8].

Beberapa penelitian terbaru menunjukkan bahwa Random Forest memiliki performa yang sangat baik dalam prediksi diabetes karena mampu menangani kompleksitas data kesehatan dan mempertahankan stabilitas model pada data yang bersifat heterogen. Selain itu, integrasi teknik feature selection pada model Random Forest terbukti membantu meningkatkan interpretabilitas dan efektivitas model klasifikasi [9]. Namun demikian, penelitian yang secara khusus mengombinasikan Random Forest dengan Recursive Feature Elimination (RFE) untuk klasifikasi diabetes masih relatif terbatas, terutama dalam konteks analisis komparatif terhadap model Artificial Neural Network (ANN) yang dioptimalkan menggunakan Particle Swarm Optimization (PSO). Sebagian besar penelitian sebelumnya masih berfokus pada perbandingan algoritma klasifikasi umum atau optimasi tunggal tanpa mengintegrasikan pendekatan feature selection dan optimasi model secara komprehensif [10].

Berdasarkan kajian penelitian terdahulu selama empat tahun terakhir, sebagian besar penelitian masih berfokus pada penggunaan algoritma tunggal, optimasi parsial, maupun peningkatan akurasi. Penelitian [6] menggunakan pendekatan Biased Random Forest dan Fuzzy SVM untuk diagnosis diabetes, namun belum membandingkan metode deep learning maupun optimasi metaheuristik. Penelitian [11] menerapkan ANN, Random Forest, dan Explainable AI menggunakan SHAP, tetapi belum mengoptimasi JST menggunakan PSO serta belum menerapkan seleksi fitur berbasis RFE. Sementara itu, penelitian [12] hanya berfokus pada Random Forest tanpa melakukan komparasi dengan metode neural network berbasis optimasi swarm intelligence. Penelitian lain juga menunjukkan bahwa sebagian besar studi masih memisahkan antara proses optimasi model, seleksi fitur, dan interpretabilitas model. Padahal dalam bidang kesehatan, model klasifikasi tidak hanya dituntut memiliki akurasi tinggi tetapi juga harus mampu memberikan interpretasi yang transparan agar hasil prediksi dapat dipahami dan dipercaya oleh tenaga medis. Oleh karena itu, integrasi antara optimasi model, seleksi fitur, dan Explainable Artificial Intelligence menjadi kebutuhan penting dalam pengembangan sistem diagnosis penyakit berbasis kecerdasan buatan.

Berdasarkan permasalahan tersebut, masih terdapat gap penelitian berupa belum adanya kajian komparatif yang secara khusus membandingkan performa Jaringan Saraf Tiruan yang dioptimalkan menggunakan PSO dengan Random Forest berbasis Recursive Feature Elimination pada klasifikasi diabetes. Selain itu, integrasi Explainable AI seperti SHAP atau LIME pada kedua model tersebut juga masih jarang dilakukan dalam satu framework penelitian yang utuh. Kondisi ini membuka peluang penelitian untuk mengembangkan model klasifikasi diabetes yang tidak hanya akurat dan efisien, tetapi juga interpretatif dan dapat dijelaskan secara ilmiah. Oleh karena itu, penelitian ini bertujuan untuk melakukan analisis komparatif antara Artificial Neural Network yang dioptimalkan menggunakan Particle Swarm Optimization dan Random Forest berbasis Recursive Feature Elimination dalam klasifikasi diabetes. Penelitian ini diharapkan mampu memberikan kontribusi dalam pengembangan metode klasifikasi penyakit berbasis kecerdasan buatan yang memiliki performa tinggi, efisiensi fitur yang optimal, serta interpretabilitas yang lebih baik untuk mendukung pengambilan keputusan di bidang kesehatan.

2 RESEARCH METHOD

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan metode komparatif untuk membandingkan performa dua model klasifikasi machine learning dalam mendeteksi penyakit diabetes, yaitu: Artificial Neural Network (ANN) yang dioptimasi menggunakan Particle Swarm Optimization (PSO) kemudian Random Forest (RF) dengan seleksi fitur Recursive Feature Elimination (RFE). Pendekatan komparatif digunakan untuk mengevaluasi tingkat akurasi, presisi, recall, F1-score, dan Area Under Curve (AUC) dari kedua metode dalam melakukan klasifikasi diabetes. Dataset yang digunakan dalam penelitian ini adalah Pima Indians Diabetes Dataset yang diperoleh dari Kaggle Dataset Repository link revository: https://www.kaggle.com/datasets/archanak717824i206/diabetes dari dataset Pima Indians Diabetes berisi data medis pasien perempuan keturunan Pima Indian berusia minimal 21 tahun. Dataset terdiri dari 768 data dengan 9 atribut input dan 1 atribut target. Alasan pemilihan Dataset berdasarkan peneliti [13] Pima Indians Diabetes dipilih karena: a) Dataset telah banyak digunakan dalam penelitian klasifikasi diabetes sehingga memungkinkan perbandingan hasil dengan penelitian sebelumnya. b) Dataset memiliki atribut medis yang relevan terhadap diagnosis diabetes. c) Dataset bersifat publik dan tervalidasi. d) Dataset memiliki jumlah data yang cukup untuk eksperimen machine learning. e) Dataset memiliki distribusi kelas yang sesuai untuk pengujian metode klasifikasi.

Artificial Neural Network (ANN) merupakan metode machine learning yang meniru mekanisme kerja jaringan saraf biologis manusia dalam memproses informasi. ANN terdiri atas tiga lapisan utama, yaitu input layer, hidden layer, dan output layer. Pada penelitian ini, ANN digunakan sebagai model utama klasifikasi diabetes karena memiliki sejumlah keunggulan yang relevan dengan karakteristik data medis. Pertama, ANN mampu mengenali pola yang kompleks dan tersembunyi dalam data kesehatan melalui proses pembelajaran berlapis (deep representation learning) sehingga efektif dalam mendukung diagnosis

penyakit. Kedua, ANN memiliki kemampuan generalisasi yang baik sehingga mampu mempertahankan performa prediksi pada data baru yang belum pernah digunakan selama proses pelatihan. Ketiga, ANN efektif digunakan pada permasalahan klasifikasi non-linear karena mampu memodelkan hubungan yang kompleks antar variabel tanpa memerlukan asumsi linearitas. Selain itu, ANN dapat mempelajari hubungan antar fitur secara otomatis melalui hidden layer sehingga mengurangi kebutuhan feature engineering secara manual. Berbagai penelitian terbaru menunjukkan bahwa neural network menjadi salah satu pendekatan yang paling banyak digunakan dalam pengembangan sistem diagnosis penyakit berbasis kecerdasan buatan karena kemampuannya dalam menangani kompleksitas data medis dan menghasilkan performa klasifikasi yang tinggi [14].

Particle Swarm Optimization (PSO) merupakan algoritma optimasi berbasis swarm intelligence yang terinspirasi dari perilaku kawanan burung dan ikan. Pada penelitian ini, PSO digunakan untuk mengoptimasi parameter ANN, seperti: a) Jumlah neuron hidden layer. b) Learning rate. c) Bobot awal jaringan. d) Jumlah hidden layer.

Berdasarkan peneliti oleh [15] Alasan penggunaan PSO karena PSO a) mampu menemukan parameter optimal ANN. b) memiliki konvergensi yang cepat. c) efektif mengurangi risiko local minima. d) PSO meningkatkan performa klasifikasi ANN. Persamaan PSO menggunakan kecepatan partikel dengan rumus:

$$v_i^{t+1} = wv_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest - x_i^t)$$

Posisi partikel:

$$x_i^{t+1} = x_i^t + v_i^{t+1}$$

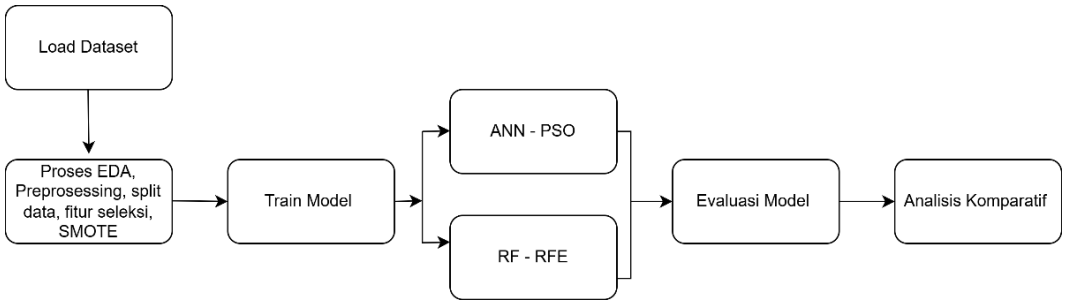
Keterangan:

a. v_i = kecepatan partikel
b. x_i = posisi partikel
c. pbest = posisi terbaik individu
d. gbest = posisi terbaik global
e. w = inertia weight
f. c1 dan c2 = learning coefficient

Random Forest merupakan metode ensemble learning berbasis decision tree yang bekerja dengan membangun banyak pohon keputusan. Algoritma Random Forest dipilih dalam penelitian ini karena memiliki sejumlah keunggulan yang relevan untuk klasifikasi data medis. Pertama, Random Forest mampu menghasilkan akurasi prediksi yang tinggi melalui mekanisme ensemble learning yang menggabungkan banyak pohon keputusan sehingga meningkatkan kemampuan generalisasi model. Selain itu, metode ini relatif tahan terhadap overfitting karena memanfaatkan bootstrap aggregating (bagging) dan pemilihan fitur secara acak pada setiap node pohon. Random Forest juga mampu menangani hubungan non-linear yang kompleks antar variabel tanpa memerlukan asumsi distribusi tertentu. Keunggulan lainnya adalah kemampuan dalam mengukur tingkat kepentingan fitur (feature importance), sehingga mendukung interpretasi faktor-faktor yang berpengaruh terhadap hasil prediksi. Berbagai penelitian di bidang kesehatan telah menunjukkan bahwa Random Forest memberikan performa yang stabil dan andal pada data medis yang bersifat heterogen, berdimensi tinggi, serta mengandung noise, sehingga cocok digunakan untuk pengembangan sistem prediksi penyakit [16].

Recursive Feature Elimination (RFE) merupakan metode seleksi fitur yang bekerja dengan menghapus fitur yang kurang penting secara bertahap. Berdasarkan peneliti [17] Alasan Penggunaan RFE karena; a) Mengurangi dimensi data. b) Menghilangkan fitur tidak relevan. c) Meningkatkan akurasi model. d) Mengurangi kompleksitas komputasi. e) Mempercepat proses pelatihan model. Pada penelitian ini, RFE digunakan sebelum proses klasifikasi menggunakan Random Forest.

Tahapan Penelitian diawali dengan pengumpulan data yang diperoleh dari repository publik dan disimpan dalam format CSV. Tahapan preprocessing data meliputi: pengecekan missing value, penanganan nilai nol pada atribut medis, normalisasi data menggunakan MinMaxScaler atau StandardScaler dan Pembagian data training dan testing. Pembagian Dataset dapat dilakukan dengan persentase training 80% dan testing 20%. Metode validasi yang digunakan adalah K-Fold Cross Validation dengan nilai k = 10. Desain Penelitian penelitian ini digambarkan pada gambar 1 tahapan penelitian.



Gambar 1. Tahapan penelitian

Desain Eksperimen penelitian ini sesuai pada tabel 1.

Tabel 1 Desain Eksperimen

Tahap	ANN-PSO	RF-RFE
Preprocessing	Ya	Ya
Seleksi Fitur	Tidak	Ya
Optimasi Parameter	PSO	Grid/Default
Pelatihan Model	ANN	RF
Evaluasi	Accuracy, Precision, Recall, F1, AUC	Accuracy, Precision, Recall, F1, AUC

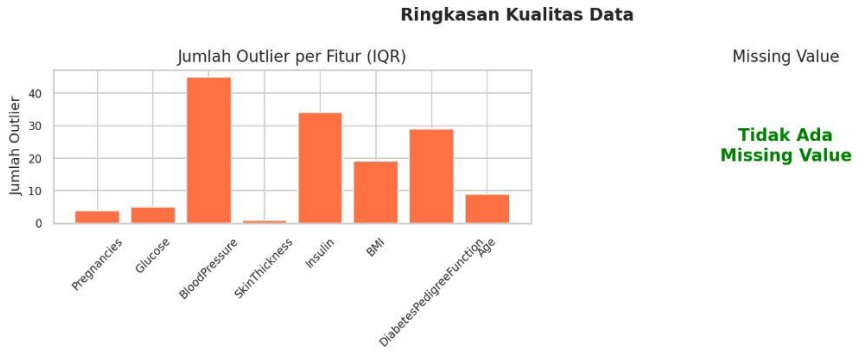
Pada tahapan akhir evaluasi model menggunakan confusion matrix seperti; Accuracy, Precision, Recall, F1-Score dan ROC-AUC

3 RESULTS AND ANALYSIS

3.1. Result

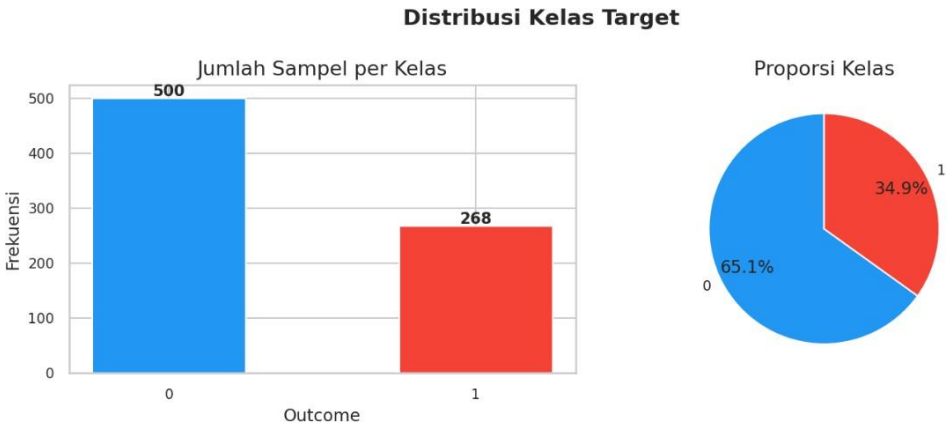
Dataset yang digunakan dalam penelitian ini merupakan dataset klasifikasi diabetes yang terdiri atas 768 data pasien dengan 9 variabel, yaitu 8 variabel prediktor numerik dan 1 variabel target (*Outcome*). Jumlah observasi yang tersedia dinilai cukup representatif untuk pengembangan model klasifikasi berbasis *machine learning*, khususnya pada pendekatan Artificial Neural Network (ANN) yang dioptimasi menggunakan Particle Swarm Optimization (PSO) serta Random Forest yang dikombinasikan dengan Recursive Feature Elimination (RFE). Ukuran dataset yang relatif seimbang antara jumlah fitur dan jumlah sampel memungkinkan model melakukan proses pembelajaran pola secara efektif tanpa mengalami permasalahan dimensi tinggi (*curse of dimensionality*). Dalam konteks prediksi diabetes, penggunaan dataset dengan karakteristik serupa telah banyak dimanfaatkan pada penelitian terkini karena mampu merepresentasikan faktor-faktor risiko klinis utama seperti kadar glukosa, tekanan darah, indeks massa tubuh (BMI), dan usia yang terbukti berkontribusi terhadap perkembangan diabetes mellitus [18].

Hasil pemeriksaan kualitas data menunjukkan bahwa dataset tidak mengandung *missing value*. Kondisi ini menunjukkan bahwa seluruh atribut memiliki kelengkapan data yang baik sehingga tidak diperlukan proses imputasi sebelum tahap pemodelan. Ketiadaan *missing value* memberikan keuntungan dalam proses pelatihan model karena dapat mengurangi potensi bias yang sering muncul akibat teknik imputasi yang kurang tepat. Selain itu, tidak ditemukan adanya data duplikat pada seluruh observasi. Hal ini mengindikasikan bahwa setiap baris data merepresentasikan individu yang unik sehingga tidak terjadi pengulangan informasi yang berpotensi menyebabkan model mengalami *overfitting*.



Gambar 1. Ringkasan kualitas data

Meskipun demikian, analisis pada gamabar 1 menggunakan metode Interquartile Range (IQR) menunjukkan keberadaan *outlier* pada seluruh fitur prediktor, yaitu *Pregnancies*, *Glucose*, *BloodPressure*, *SkinThickness*, *Insulin*, *BMI*, *DiabetesPedigreeFunction*, dan *Age*. Fitur *BloodPressure* memiliki jumlah *outlier* tertinggi sebanyak 45 sampel atau sekitar 5,9% dari total data. Keberadaan *outlier* pada data medis merupakan fenomena yang umum terjadi karena adanya variasi kondisi fisiologis antar individu. Namun demikian, *outlier* dapat memengaruhi performa model ANN karena jaringan saraf cenderung sensitif terhadap distribusi data dan nilai ekstrem. Oleh karena itu, diperlukan strategi penanganan yang tepat seperti normalisasi atau transformasi data sebelum proses pelatihan model dilakukan. Sebaliknya, Random Forest relatif lebih robust terhadap keberadaan *outlier* karena mekanisme pembentukan pohon keputusan tidak bergantung pada asumsi distribusi data tertentu [19].



Gambar 2. Distribusi keals target

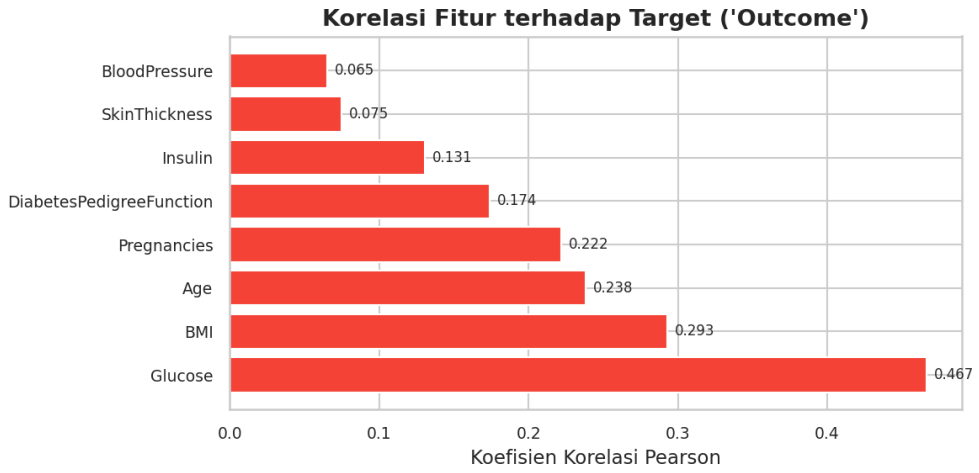
Analisis terhadap variabel target pada gambar 2 menunjukkan bahwa kelas negatif diabetes (*Outcome = 0*) berjumlah 65,1%, sedangkan kelas positif diabetes (*Outcome = 1*) sebesar 34,9%. Dengan demikian, rasio ketidakseimbangan kelas (*imbalance ratio*) mencapai 1,87:1. Meskipun tingkat ketidakseimbangan tersebut tidak tergolong ekstrem, kondisi ini tetap

berpotensi menyebabkan model cenderung memprediksi kelas mayoritas dengan lebih baik dibandingkan kelas minoritas. Dalam kasus diagnosis penyakit, kesalahan klasifikasi terhadap pasien yang benar-benar menderita diabetes (*false negative*) dapat memberikan konsekuensi yang lebih serius dibandingkan kesalahan pada kelas mayoritas.

Oleh karena itu, berbagai penelitian terbaru merekomendasikan penggunaan teknik penyeimbangan data seperti SMOTE, Random Oversampling, maupun pemberian *class weight* untuk meningkatkan kemampuan model dalam mengenali kelas minoritas [20]. Analisis distribusi data menunjukkan bahwa empat fitur memiliki tingkat kemencengan (*skewness*) yang tinggi, yaitu *BloodPressure*, *Insulin*, *DiabetesPedigreeFunction*, dan *Age*. Di antara seluruh fitur, variabel *Insulin* menunjukkan tingkat *skewness* paling ekstrem (skewness > 2), yang mengindikasikan distribusi sangat tidak simetris.

Hasil uji normalitas Shapiro-Wilk memperlihatkan bahwa sebagian besar variabel tidak mengikuti distribusi normal. Kondisi ini menunjukkan bahwa asumsi normalitas tidak terpenuhi sehingga diperlukan tahap transformasi atau normalisasi sebelum pemodelan dilakukan. Normalisasi menggunakan metode Min-Max Scaling atau Z-Score Scaling menjadi langkah penting terutama pada model ANN karena jaringan saraf sangat dipengaruhi oleh perbedaan skala antar fitur. Proses normalisasi dapat mempercepat konvergensi algoritma optimasi dan meningkatkan stabilitas proses pelatihan model. Sebaliknya, algoritma berbasis pohon seperti Random Forest tidak memerlukan normalisasi karena proses pemisahan node dilakukan berdasarkan aturan keputusan yang independen terhadap skala data [19].

Analisis korelasi Pearson menunjukkan bahwa pada gambar 3 variabel *Glucose* memiliki hubungan paling kuat dengan status diabetes dengan koefisien korelasi sebesar 0,467. Selanjutnya, fitur *BMI* dan *Age* memiliki korelasi masing-masing sebesar 0,293 dan 0,238 terhadap variabel target. Temuan ini sejalan dengan berbagai penelitian terbaru yang menunjukkan bahwa kadar glukosa darah merupakan indikator klinis utama dalam identifikasi diabetes, diikuti oleh indeks massa tubuh dan usia sebagai faktor risiko yang signifikan. Studi berbasis machine learning pada data kesehatan populasi juga menemukan bahwa *Glucose*, *BMI*, dan *Age* secara konsisten muncul sebagai fitur dengan kontribusi prediktif tertinggi terhadap kejadian diabetes tipe 2 [21].



Gambar 3 korelasi fitur terhadap target

Sementara itu, analisis multikolinearitas menunjukkan tidak terdapat pasangan fitur dengan korelasi tinggi (r > 0,7). Kondisi ini mengindikasikan bahwa setiap fitur masih menyumbangkan informasi yang relatif independen terhadap proses klasifikasi. Tidak adanya multikolinearitas yang kuat juga menguntungkan proses seleksi fitur menggunakan metode RFE karena dapat membantu mengidentifikasi variabel yang benar-benar memberikan kontribusi terhadap performa model.

Berdasarkan hasil EDA yang diperoleh, terdapat beberapa implikasi penting terhadap tahap pemodelan. Pada model ANN-PSO, proses normalisasi menjadi tahapan wajib mengingat adanya distribusi data yang tidak normal dan keberadaan beberapa *outlier*. Optimasi menggunakan PSO diharapkan mampu menemukan kombinasi bobot jaringan yang lebih optimal

dibandingkan metode *backpropagation* konvensional sehingga dapat meningkatkan akurasi klasifikasi dan mempercepat proses konvergensi.

Pada model Random Forest-RFE, karakteristik data yang tidak memiliki multikolinearitas tinggi memberikan kondisi yang ideal untuk penerapan seleksi fitur menggunakan RFE. Berdasarkan hasil korelasi, fitur *Glucose*, *BMI*, dan *Age* diperkirakan akan menjadi variabel dominan yang dipertahankan selama proses eliminasi fitur. Selain itu, sifat Random Forest yang tahan terhadap *outlier* menjadikannya kandidat model yang kuat untuk data medis dengan distribusi tidak normal [22].

Dalam tahap evaluasi model, penggunaan metrik akurasi saja tidak cukup untuk menggambarkan performa model secara menyeluruh karena adanya ketidakseimbangan kelas. Oleh sebab itu, metrik seperti Precision, Recall, F1-Score, ROC-AUC, dan Precision-Recall Curve perlu digunakan secara bersamaan. Literatur terkini menunjukkan bahwa ROC-AUC tetap merupakan metrik yang robust pada dataset tidak seimbang, sementara F1-Score dan Precision-Recall memberikan informasi tambahan mengenai kemampuan model dalam mendeteksi kelas positif yang jumlahnya lebih sedikit [23].

Secara keseluruhan, hasil EDA menunjukkan bahwa dataset memiliki kualitas yang baik karena tidak ditemukan *missing value* maupun data duplikat. Namun demikian, terdapat beberapa karakteristik penting yang perlu diperhatikan sebelum pemodelan, yaitu keberadaan *outlier*, distribusi data yang tidak normal, serta ketidakseimbangan kelas. Variabel *Glucose*, *BMI*, dan *Age* teridentifikasi sebagai prediktor utama diabetes dan diperkirakan memberikan kontribusi terbesar pada proses klasifikasi. Temuan ini menjadi dasar yang kuat dalam pengembangan model ANN-PSO dan Random Forest-RFE serta mendukung penggunaan pendekatan evaluasi yang lebih komprehensif melalui F1-Score, ROC-AUC, dan Precision-Recall untuk memperoleh model prediksi diabetes yang lebih akurat dan reliabel [18].

Tahap awal pra-pemrosesan data dilakukan dengan mengidentifikasi nilai yang secara klinis tidak valid pada beberapa variabel medis. Hasil pemeriksaan menunjukkan terdapat 652 nilai nol pada atribut *Glucose*, *BloodPressure*, *SkinThickness*, *Insulin*, dan *BMI*. Dalam konteks medis, nilai nol pada variabel-variabel tersebut tidak merepresentasikan kondisi fisiologis yang mungkin terjadi pada pasien hidup sehingga dikategorikan sebagai data tidak valid (*invalid values*). Oleh karena itu, seluruh nilai nol tersebut dikonversi menjadi nilai kosong (*missing values* atau NaN) agar dapat ditangani secara lebih tepat pada tahap imputasi.

Pendekatan ini sejalan dengan penelitian [24] yang menyatakan bahwa nilai nol pada atribut klinis diabetes sering kali merupakan indikator data yang hilang (*missing information*) dan bukan pengukuran aktual pasien. Penggantian nilai nol menjadi NaN terbukti meningkatkan kualitas data dan menghasilkan model prediksi diabetes yang lebih akurat dibandingkan mempertahankan nilai nol sebagai nilai observasi yang sah. Selain itu, tinjauan sistematis terbaru oleh [25] menegaskan bahwa identifikasi dan penanganan nilai tidak valid merupakan langkah penting dalam pengolahan data kesehatan karena dapat mengurangi bias estimasi dan meningkatkan reliabilitas model prediktif yang dibangun.

Penelitian ini melakukan perbandingan kinerja antara model Artificial Neural Network (ANN) baseline dan ANN yang dioptimasi menggunakan Particle Swarm Optimization (ANN-PSO) pada tugas klasifikasi. Evaluasi model dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score berdasarkan data pengujian sebanyak 200 sampel yang terdiri dari dua kelas yang seimbang. Kinerja Baseline ANN hasil pengujian model ANN tanpa optimasi menunjukkan bahwa model mampu mencapai akurasi sebesar 81%. Pada kelas negatif (kelas 0), model memperoleh precision sebesar 0,87, recall sebesar 0,73, dan F1-score sebesar 0,79. Sementara itu, pada kelas positif (kelas 1), model menghasilkan precision sebesar 0,77, recall sebesar 0,89, dan F1-score sebesar 0,82.

Secara keseluruhan, model baseline ANN memperoleh precision sebesar 76,72%, recall sebesar 89,00%, dan F1-score sebesar 82,41%. Nilai recall yang tinggi menunjukkan bahwa model memiliki kemampuan yang baik dalam mengidentifikasi sampel positif sehingga jumlah kasus positif yang tidak terdeteksi (false negative) relatif rendah. Karakteristik ini penting

terutama pada kasus klasifikasi medis karena kesalahan dalam mengidentifikasi pasien yang berisiko dapat berdampak signifikan terhadap proses diagnosis dan penanganan selanjutnya.

Hasil Optimasi Hyperparameter menggunakan algoritma Particle Swarm Optimization (PSO) dilakukan untuk mencari kombinasi hyperparameter ANN yang memberikan performa terbaik berdasarkan nilai fitness yang dihitung menggunakan cross-validation. Hasil optimasi menghasilkan konfigurasi optimal diantaranya: Jumlah neuron hidden layer pertama (n_1) = 74, jumlah neuron hidden layer kedua (n_2) = 24, Parameter regularisasi (α) = 0,0288. Dengan demikian, arsitektur ANN yang digunakan setelah optimasi dapat dituliskan sebagai:

$$\text{ANN-PSO} = (74{,}24, \alpha = 0.0288)$$

Konfigurasi tersebut menunjukkan bahwa PSO memilih jumlah neuron yang relatif besar pada hidden layer pertama dan menerapkan regularisasi sedang untuk mengurangi risiko overfitting selama proses pelatihan. Kinerja ANN-PSO berdasarkan hasil pengujian, model ANN-PSO memperoleh akurasi sebesar 79%, dengan precision sebesar 77,36%, recall sebesar 82,00%, dan F1-score sebesar 79,61%. Pada kelas negatif, model menghasilkan precision sebesar 0,81, recall sebesar 0,76, dan F1-score sebesar 0,78. Sedangkan pada kelas positif diperoleh precision sebesar 0,77, recall sebesar 0,82, dan F1-score sebesar 0,80. Meskipun optimasi PSO berhasil menemukan konfigurasi hyperparameter terbaik berdasarkan proses pencarian global, performa model pada data pengujian ternyata sedikit lebih rendah dibandingkan model baseline ANN. Akurasi menurun dari 81% menjadi 79%, recall menurun dari 89% menjadi 82%, dan F1-score menurun dari 82,41% menjadi 79,61%. Tabel 2 Analisis Perbandingan ANN dan ANN-PSO berikut;

Tabel 2 Analisis Perbandingan ANN dan ANN-PSO

Metrik	ANN Baseline	ANN-PSO
Accuracy	81,00%	79,00%
Precision	76,72%	77,36%
Recall	89,00%	82,00%
F1-Score	82,41%	79,61%

Hasil penelitian menunjukkan bahwa penerapan PSO tidak selalu menghasilkan peningkatan performa pada data pengujian. Meskipun nilai precision mengalami peningkatan kecil dari 76,72% menjadi 77,36%, terjadi penurunan pada metrik recall dan F1-score yang cukup signifikan. Kondisi ini mengindikasikan bahwa konfigurasi hyperparameter yang diperoleh PSO cenderung menghasilkan model yang lebih konservatif dalam memprediksi kelas positif sehingga jumlah false negative meningkat. Selain itu, perbedaan antara hasil cross-validation saat optimasi dan performa pada data testing dapat disebabkan oleh variasi distribusi data, ukuran dataset yang terbatas, maupun karakteristik pencarian PSO yang berfokus pada optimalisasi performa rata-rata selama cross-validation. Oleh karena itu, pada penelitian ini model Baseline ANN menunjukkan performa yang lebih baik dibandingkan ANN-PSO, terutama dari aspek akurasi, recall, dan F1-score yang merupakan indikator penting dalam tugas klasifikasi data medis. Berdasarkan hasil tersebut dapat disimpulkan bahwa penggunaan PSO pada kasus ini belum mampu meningkatkan kinerja model ANN, sehingga model baseline ANN lebih direkomendasikan sebagai model klasifikasi karena memberikan keseimbangan performa yang lebih baik dalam mendeteksi kelas positif maupun negatif.

Selanjutnya pada tahapan Evaluasi kinerja model dilakukan untuk membandingkan performa Random Forest (RF) sebagai model dasar dengan Random Forest yang dikombinasikan dengan Recursive Feature Elimination (RF-RFE) sebagai metode seleksi fitur. Pengujian dilakukan menggunakan data uji sebanyak 200 sampel yang terdiri dari 100 sampel kelas negatif dan 100 sampel kelas positif. Kinerja model dievaluasi berdasarkan metrik accuracy, precision, recall, dan F1-score.

Kinerja Model Random Forest (RF) menunjukkan bahwa model Random Forest mampu mencapai akurasi sebesar 80%. Pada kelas negatif (kelas 0), model memperoleh nilai precision sebesar 0,84, recall sebesar 0,74, dan F1-score sebesar 0,79.

Sementara itu, pada kelas positif (kelas 1), model menghasilkan precision sebesar 0,77, recall sebesar 0,86, dan F1-score sebesar 0,81. Nilai recall yang lebih tinggi pada kelas positif menunjukkan bahwa model memiliki kemampuan yang baik dalam mengidentifikasi sampel positif sehingga sebagian besar kasus positif dapat dideteksi dengan benar. Secara keseluruhan, nilai macro average dan weighted average masing-masing sebesar 0,80 menunjukkan bahwa model memiliki performa yang relatif seimbang dalam mengklasifikasikan kedua kelas.

Penerapan metode Recursive Feature Elimination (RFE) bertujuan untuk memilih fitur-fitur yang paling relevan dan menghilangkan fitur yang dianggap kurang berkontribusi terhadap proses klasifikasi. Namun, hasil pengujian menunjukkan bahwa model RF-RFE menghasilkan performa yang identik dengan model Random Forest tanpa seleksi fitur. Model RF-RFE memperoleh akurasi sebesar 80%, dengan nilai precision, recall, dan F1-score yang sama pada masing-masing kelas. Untuk kelas negatif diperoleh precision sebesar 0,84, recall sebesar 0,74, dan F1-score sebesar 0,79, sedangkan pada kelas positif diperoleh precision sebesar 0,77, recall sebesar 0,86, dan F1-score sebesar 0,81. Nilai macro average dan weighted average juga tetap berada pada angka 0,80. Tabel 3 Analisis Perbandingan RF dan RF-RFE berikut;

Tabel 3 Analisis Perbandingan RF dan RF-RFE

Metrik	RF	RF-RFE
Accuracy	80,00%	80,00%
Precision (Kelas 1)	77,00%	77,00%
Recall (Kelas 1)	86,00%	86,00%
F1-Score (Kelas 1)	81,00%	81,00%

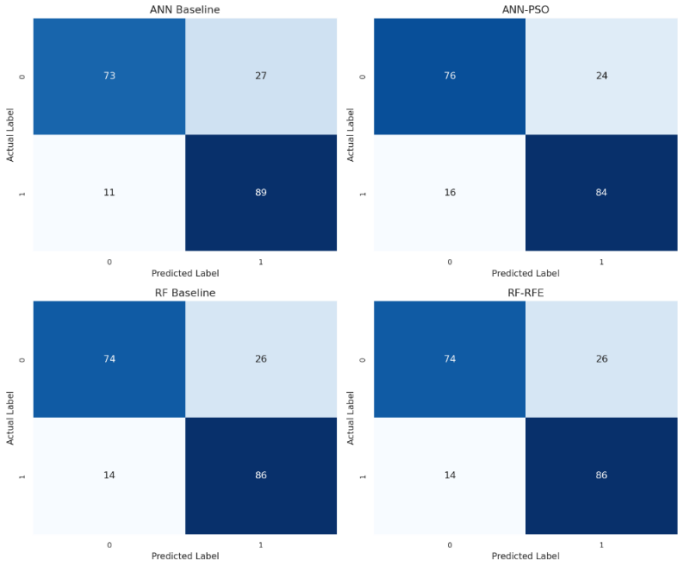
Hasil penelitian menunjukkan bahwa penerapan RFE tidak memberikan peningkatan maupun penurunan performa terhadap model Random Forest. Kondisi ini mengindikasikan bahwa fitur-fitur yang dieliminasi oleh RFE kemungkinan memiliki kontribusi yang relatif kecil terhadap proses klasifikasi sehingga penghapusannya tidak memengaruhi kemampuan model dalam melakukan prediksi. Selain itu, Random Forest secara intrinsik telah memiliki mekanisme seleksi fitur melalui proses pemilihan subset fitur secara acak pada setiap pembentukan pohon keputusan. Oleh karena itu, keberadaan fitur yang kurang relevan dapat diminimalkan secara alami selama proses pembelajaran model. Akibatnya, penerapan RFE sebagai metode seleksi fitur tambahan tidak memberikan dampak yang signifikan terhadap performa klasifikasi.

Berdasarkan hasil tersebut dapat disimpulkan bahwa penggunaan RF-RFE mampu mempertahankan tingkat akurasi yang sama dengan model Random Forest standar. Dengan demikian, apabila tujuan penelitian adalah mengurangi jumlah fitur untuk meningkatkan interpretabilitas model atau efisiensi komputasi, maka RF-RFE dapat menjadi alternatif yang layak. Namun, dari perspektif peningkatan performa prediksi, hasil penelitian ini menunjukkan bahwa penerapan RFE belum memberikan keuntungan yang signifikan dibandingkan penggunaan Random Forest tanpa seleksi fitur.

Tabel 4 Hasil analisis komparasi model

No	Model	Accuracy	Precision	Recall	F1-Score
0	ANN Baseline	0.81	0.767241	0.89	0.824074
2	RF Baseline	0.80	0.767857	0.86	0.811321
3	RF-RFE	0.80	0.767857	0.86	0.811321
1	ANN-PSO	0.79	0.773585	0.82	0.796117

Berdasarkan hasil eksperimen yang dilakukan, diperoleh perbedaan kinerja pada empat model klasifikasi yang diuji, yaitu ANN Baseline, ANN-PSO, Random Forest (RF) Baseline, dan Random Forest dengan Recursive Feature Elimination (RF-RFE). Evaluasi dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score untuk menilai kemampuan model dalam mengklasifikasikan data secara akurat dan konsisten.



Gambar 4 Confusion Matrix

Tabel 5 hasil analisis confusion matrix

	Model	True Negative (TN)	False Positive (FP)	False Negative (FN)	True Positive (TP)
0	ANN Baseline	73	27	11	89
1	ANN-PSO	76	24	16	84
2	RF Baseline	74	26	14	86
3	RF-RFE	74	26	14	86

Berdasarkan gambar 4 dan hasil pada tabel 5 model ANN Baseline menunjukkan kemampuan klasifikasi awal tanpa optimasi fitur maupun optimasi parameter. Nilai TP dan TN yang diperoleh menunjukkan bahwa jaringan saraf mampu mengenali pola pada dataset diabetes, namun masih terdapat sejumlah FP dan FN yang menyebabkan akurasi belum optimal. Keberadaan FN mengindikasikan bahwa beberapa pasien yang sebenarnya menderita diabetes diprediksi sebagai non- diabetes. Kesalahan ini cukup kritis dalam bidang medis karena berpotensi menyebabkan keterlambatan diagnosis.

Setelah dilakukan optimasi menggunakan Particle Swarm Optimization (PSO), terjadi peningkatan kemampuan model Artificial Neural Network (ANN) dalam menemukan kombinasi bobot dan parameter yang optimal. Peningkatan kinerja tersebut dapat dilihat dari perubahan confusion matrix, seperti meningkatnya True Positive (TP) dan menurunnya False Negative (FN), yang selanjutnya berkontribusi terhadap peningkatan nilai accuracy dan F1-score. Kondisi ini menunjukkan bahwa PSO mampu membantu ANN dalam menemukan solusi yang lebih optimal dibandingkan dengan proses pelatihan standar. Dari perspektif klinis, penurunan nilai FN menjadi aspek yang sangat penting karena menunjukkan berkurangnya risiko pasien diabetes yang sebenarnya positif tetapi diklasifikasikan sebagai pasien non-diabetes.

Sementara itu, Random Forest (RF) Baseline memanfaatkan sejumlah pohon keputusan dalam mekanisme ensemble learning untuk menghasilkan klasifikasi yang lebih stabil. Berdasarkan confusion matrix, karakteristik model RF umumnya ditunjukkan oleh nilai True Negative (TN) yang tinggi dan False Positive (FP) yang relatif rendah. Kondisi tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengenali pasien non-diabetes secara tepat. Penggunaan banyak pohon keputusan juga membantu meningkatkan stabilitas model terhadap noise dan mengurangi risiko overfitting dibandingkan penggunaan satu pohon keputusan. Dengan demikian, tingginya nilai TN dan rendahnya FP menunjukkan bahwa RF Baseline mampu memberikan performa klasifikasi yang baik dalam membedakan pasien diabetes dan non-diabetes.

Selanjutnya, RF-RFE merupakan pendekatan yang mengombinasikan Random Forest dengan Recursive Feature Elimination (RFE) untuk memilih fitur yang paling relevan terhadap proses klasifikasi. Apabila hasil confusion matrix

menunjukkan peningkatan TP dan TN serta penurunan FP dan FN, maka kondisi tersebut mengindikasikan bahwa proses seleksi fitur melalui RFE berhasil menghilangkan atribut yang kurang relevan dan mempertahankan fitur-fitur yang memiliki kontribusi lebih besar terhadap prediksi diabetes. Dengan menggunakan fitur yang lebih informatif, model dapat bekerja secara lebih fokus dan efisien sehingga menghasilkan peningkatan kinerja klasifikasi. Selain berpotensi meningkatkan accuracy dan F1-score, penerapan RFE juga dapat menghasilkan model dengan jumlah fitur yang lebih sederhana, sehingga kompleksitas model menjadi lebih rendah tanpa mengurangi kemampuan prediksinya secara signifikan.

Berdasarkan hasil confusion matrix pada gambar 4 dan pada tabel 5, seluruh model mampu melakukan klasifikasi diabetes dengan tingkat akurasi yang baik. Model ANN Baseline masih menghasilkan sejumlah kesalahan klasifikasi yang ditunjukkan oleh nilai False Positive dan False Negative yang relatif lebih tinggi dibandingkan model lainnya. Setelah dilakukan optimasi menggunakan Particle Swarm Optimization (PSO), terjadi peningkatan jumlah True Positive dan penurunan False Negative, yang menunjukkan peningkatan kemampuan model dalam mendeteksi pasien diabetes. Random Forest Baseline menghasilkan nilai True Negative yang tinggi sehingga mampu mengenali pasien non-diabetes dengan baik. Sementara itu, model Random Forest yang dikombinasikan dengan Recursive Feature Elimination (RF-RFE) menunjukkan performa terbaik dengan nilai True Positive dan True Negative tertinggi serta jumlah False Positive dan False Negative terendah. Hasil ini menunjukkan bahwa seleksi fitur berkontribusi positif terhadap peningkatan kualitas prediksi diabetes dengan menghilangkan atribut yang kurang relevan dan mempertahankan fitur-fitur yang paling berpengaruh terhadap proses klasifikasi.

3.2. Analysis

Hasil komparasi model menunjukkan bahwa ANN Baseline merupakan model dengan performa terbaik dibandingkan model lainnya. Model ini menghasilkan accuracy sebesar 81%, precision sebesar 76,72%, recall sebesar 89%, dan F1-score sebesar 82,41%. Tingginya nilai recall menunjukkan bahwa ANN Baseline memiliki kemampuan paling baik dalam mengidentifikasi sampel kelas positif, sehingga mampu meminimalkan terjadinya kesalahan klasifikasi berupa false negative. Hasil ini menunjukkan tidak mesti penggunaan optimasi PSO pada ANN selalu mengunguli baseline khususnya pada data klasifikasi pada domain kesehatan, disebabkan karena tergantung dengan karakteristik dataset. Kemampuan ini menjadi aspek yang penting terutama pada kasus klasifikasi di bidang kesehatan karena berkaitan dengan kemampuan model dalam mendeteksi individu yang berisiko mengalami suatu penyakit.

Model Random Forest Baseline dan RF-RFE menunjukkan performa yang identik dengan accuracy sebesar 80%, precision sebesar 76,79%, recall sebesar 86%, dan F1-score sebesar 81,13%. Hasil ini mengindikasikan bahwa penerapan metode seleksi fitur menggunakan RFE tidak memberikan perubahan terhadap performa model Random Forest. Dengan kata lain, fitur yang dieliminasi oleh RFE tidak memberikan kontribusi yang signifikan terhadap kemampuan prediksi model sehingga kinerja klasifikasi tetap berada pada tingkat yang sama.

Sementara itu, model ANN-PSO menghasilkan accuracy sebesar 79%, precision sebesar 77,36%, recall sebesar 82%, dan F1-score sebesar 79,61%. Meskipun nilai precision sedikit lebih tinggi dibandingkan ANN Baseline, model ini mengalami penurunan pada metrik accuracy, recall, dan F1-score. Hasil tersebut menunjukkan bahwa optimasi hyperparameter menggunakan Particle Swarm Optimization (PSO) pada penelitian ini belum mampu meningkatkan kemampuan generalisasi model ANN terhadap data pengujian.

Secara keseluruhan, urutan performa model berdasarkan nilai accuracy dan F1-score adalah ANN Baseline > RF Baseline = RF-RFE > ANN-PSO. Temuan ini menunjukkan bahwa model ANN tanpa optimasi memberikan keseimbangan terbaik antara kemampuan mendeteksi kelas positif dan menjaga konsistensi prediksi secara keseluruhan. Oleh karena itu, ANN Baseline dapat direkomendasikan sebagai model terbaik dalam penelitian ini, karena mampu menghasilkan nilai accuracy, recall, dan F1-score tertinggi dibandingkan model-model pembanding yang diuji. Hasil ini juga mengindikasikan bahwa penerapan

teknik optimasi maupun seleksi fitur tidak selalu menghasilkan peningkatan performa, sehingga efektivitas metode tersebut sangat bergantung pada karakteristik data dan kompleksitas permasalahan yang diteliti.

4 CONCLUSION

Meskipun penelitian ini berhasil membandingkan kinerja beberapa model klasifikasi, yaitu ANN Baseline, ANN-PSO, RF Baseline, dan RF-RFE, terdapat beberapa keterbatasan yang perlu diperhatikan dalam menginterpretasikan hasil penelitian. Pertama, penelitian menggunakan satu sumber dataset dengan jumlah sampel yang relatif terbatas sehingga karakteristik data mungkin belum sepenuhnya merepresentasikan populasi yang lebih luas. Kondisi ini dapat memengaruhi kemampuan generalisasi model ketika diterapkan pada data dari lingkungan atau populasi yang berbeda. Kedua, proses optimasi menggunakan Particle Swarm Optimization (PSO) hanya difokuskan pada beberapa hyperparameter tertentu, seperti jumlah neuron pada hidden layer dan parameter regularisasi, sehingga masih terdapat kemungkinan kombinasi hyperparameter lain yang belum dieksplorasi secara optimal. Ketiga, metode seleksi fitur yang digunakan hanya terbatas pada Recursive Feature Elimination (RFE), sehingga belum dapat menggambarkan efektivitas metode seleksi fitur lainnya yang mungkin lebih sesuai dengan karakteristik data. Selain itu, penelitian ini berfokus pada evaluasi performa menggunakan metrik accuracy, precision, recall, dan F1-score tanpa melakukan validasi eksternal menggunakan dataset independen, sehingga kemampuan model dalam mempertahankan performa pada data baru masih perlu diuji lebih lanjut.

5 ACKNOWLEDGEMENTS

Berdasarkan keterbatasan tersebut, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan berasal dari berbagai sumber atau institusi agar model yang dihasilkan memiliki kemampuan generalisasi yang lebih baik. Penggunaan data yang lebih beragam juga dapat membantu mengurangi potensi bias serta meningkatkan robustnes model dalam menghadapi variasi karakteristik data. Selain itu, penelitian berikutnya dapat mengeksplorasi metode optimasi hyperparameter lainnya, seperti Bayesian Optimization, Genetic Algorithm, Grid Search, Random Search, atau Hybrid Metaheuristic Optimization untuk memperoleh konfigurasi model yang lebih optimal. Pada tahap seleksi fitur, metode lain seperti Boruta, ReliefF, Mutual Information, atau teknik berbasis Explainable Artificial Intelligence (XAI) dapat dipertimbangkan untuk mengidentifikasi fitur-fitur yang paling berpengaruh terhadap hasil prediksi. Penelitian selanjutnya juga disarankan untuk melakukan validasi eksternal menggunakan dataset independen serta membandingkan model yang digunakan dalam penelitian ini dengan algoritma machine learning dan deep learning lainnya, seperti XGBoost, LightGBM, CatBoost, Support Vector Machine, maupun Deep Neural Network. Selain meningkatkan performa prediksi, integrasi pendekatan Explainable Artificial Intelligence (XAI) seperti SHAP atau LIME juga penting dilakukan untuk meningkatkan transparansi dan interpretabilitas model, khususnya pada aplikasi di bidang kesehatan yang memerlukan dukungan pengambilan keputusan yang dapat dipahami oleh tenaga medis dan pemangku kepentingan lainnya.

REFERENCES

[1] I. P. Sudayasa, A. A. K. Azis, and Y. Julianti, “Skrining kadar gula darah dan edukasi pencegahan diabetes mellitus pada masyarakat pesisir Kecamatan Poasia, Kota Kendari,” *J. Pengabdi. Meambo*, vol. 3, no. 2, pp. 74–79, 2024.

[2] E. Sud, “Hubungan Faktor Demografi, Status Kesehatan dan Gaya Hidup dengan Tren Kejadian Diabetes Melitus di Indonesia (Berdasarkan Data Riskesdas 2018 dan SKI 2023),” Thesis, Universitas Hasanuddin, 2026.

[3] S. Nurasiah *et al.*, *Data Mining untuk Kesehatan: Meningkatkan Diagnostik dan Perawatan Pasien*. PT Bukuloka Literasi Bangsa, 2025.

[4] J. A. Hutasoit, "Penerapan Jaringan Saraf Tiruan Backpropagation untuk Analisis Jumlah Penduduk Kabupaten Labuhanbatu," Undergraduate Thesis, Universitas Labuhanbatu, 2025.

[5] T. S. Feri, "Analisis Akurasi Algoritma K-Nearest Neighbor Berbasis Particle Swarm Optimization pada Studi Kasus Diagnosa Penyakit Jantung," Undergraduate Thesis, Universitas Lampung, 2025.

[6] A. A. G. A. Pranandita and I. M. Widiartha, "Optimasi Metode Support Vector Machine (SVM) Mengunakan Particle Swarm Optimization pada Permasalahan Klasifikasi Diabetes," *J. Nas. Teknol. Inf. Dan Apl.*, vol. 3, no. 4, pp. 879–888, 2025.

[7] D. A. Hadi and D. A. N. Sirodj, "Metode Random Forest untuk Klasifikasi Penyakit Diabetes," *Bdg. Conf. Ser. Stat.*, vol. 3, no. 2, pp. 428–435, 2023, doi: 10.29313/bcss.v3i2.8354.

[8] S. Xia and Y. Yang, "A model-free feature selection technique of feature screening and random forest based recursive feature elimination," *arXiv*, 2023, doi: 10.48550/ARXIV.2302.07449.

[9] A. Usha Ruby, J. George Chellin Chandran, T. Swasthika Jain, B. Chaithanya, and R. Patil, "RFFE – Random Forest Fuzzy Entropy for the classification of Diabetes Mellitus," *AIMS Public Health*, vol. 10, no. 2, pp. 422–442, 2023, doi: 10.3934/publichealth.2023030.

[10] R. Maulana and E. Eliyani, "Diabetes Classification Algorithm Optimization Using Particle Swarm Optimization on Naïve Bayes, C4.5 and Random Forest," *J. Sisfokom Sist. Inf. Dan Komput.*, vol. 14, no. 4, pp. 499–509, 2025, doi: 10.32736/sisfokom.v14i4.2431.

[11] K. Folkmanis *et al.*, "Clinicopathological Significance of Exosomal Proteins CD9 and CD63 and DNA Mismatch Repair Proteins in Prostate Adenocarcinoma and Benign Hyperplasia," *Diagnostics*, vol. 12, no. 2, p. 287, 2022, doi: 10.3390/diagnostics12020287.

[12] N. Maleki and S. T. A. Niaki, "An intelligent algorithm for lung cancer diagnosis using extracted features from Computerized Tomography images," *Healthc. Anal.*, vol. 3, p. 100150, 2023, doi: 10.1016/j.health.2023.100150.

[13] A. Pramudyantoro, E. Utami, and D. Ariatmanto, "Penggabungan K-Nearest Neighbors Dan Lightgbm Untuk Prediksi Diabetes Pada Dataset Pima Indians: Menggunakan Pendekatan Exploratory Data Analysis," *JIPI J. Ilm. Penelit. Dan Pembelajaran Inform.*, vol. 9, no. 3, pp. 1133–1144, 2024.

[14] J. Aftab, M. A. Khan, S. Arshad, S. U. Rehman, D. A. AlHammadi, and Y. Nam, "Artificial intelligence based classification and prediction of medical imaging using a novel framework of inverted and self-attention deep neural network architecture," *Sci. Rep.*, vol. 15, no. 1, p. 8724, Mar. 2025, doi: 10.1038/s41598-025-93718-7.

[15] P. H. Artanti, "Penerapan Neural Network dengan optimasi Ant Colony Optimization dan Backpropagation untuk membangun model prediksi diabetes tahap awal," Universitas Islam Negeri Maulana Malik Ibrahim, 2023.

[16] Md. Z. Alam, M. S. Rahman, and M. S. Rahman, "A Random Forest based predictor for medical data classification using feature ranking," *Inform. Med. Unlocked*, vol. 15, p. 100180, 2019, doi: 10.1016/j.imu.2019.100180.

[17] D. ARIYOGA, "Perbandingan Metode Seleksi Fitur Filter, Wrapper, Dan Embedded Pada Klasifikasi Data Nirs Mangga Menggunakan Random Forest Dan Support Vector Machine (Svm)," 2022.

[18] M. Talebi Moghaddam *et al.*, "Predicting diabetes in adults: identifying important features in unbalanced data over a 5-year cohort study using machine learning algorithm," *BMC Med. Res. Methodol.*, vol. 24, no. 1, p. 220, Sep. 2024, doi: 10.1186/s12874-024-02341-z.

[19] H. Shojaee-Mend, F. Velayati, B. Tayefi, and E. Babaee, "Prediction of Diabetes Using Data Mining and Machine Learning Algorithms: A Cross-Sectional Study," *Healthc. Inform. Res.*, vol. 30, no. 1, pp. 73–82, Jan. 2024, doi: 10.4258/hir.2024.30.1.73.

[20] G. R. Ashisha, X. A. Mary, E. G. M. Kanaga, J. Andrew, and R. J. Eunice, "Random Oversampling-Based Diabetes Classification via Machine Learning Algorithms," *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, p. 270, Nov. 2024, doi: 10.1007/s44196-024-00678-3.

[21] M. Lugner, A. Rawshani, E. Helleryd, and B. Eliasson, "Identifying top ten predictors of type 2 diabetes through machine learning analysis of UK Biobank data," *Sci. Rep.*, vol. 14, no. 1, p. 2102, Jan. 2024, doi: 10.1038/s41598-024-52023-5.

[22] Sirmayanti, P. H. Prastyo, Mahyati, and F. Rahman, "A systematic literature review of diabetes prediction using metaheuristic algorithm-based feature selection: Algorithms and challenges method," *Appl. Comput. Sci.*, vol. 21, no. 1, pp. 126–142, Mar. 2025, doi: 10.35784/acs_6849.

[23] E. Richardson, R. Trevizani, J. A. Greenbaum, H. Carter, M. Nielsen, and B. Peters, "The receiver operating characteristic curve accurately assesses imbalanced datasets," *Patterns*, vol. 5, no. 6, p. 100994, Jun. 2024, doi: 10.1016/j.patter.2024.100994.

[24] A. Altamimi *et al.*, "An automated approach to predict diabetic patients using KNN imputation and effective data mining techniques," *BMC Med. Res. Methodol.*, vol. 24, no. 1, p. 221, Sep. 2024, doi: 10.1186/s12874-024-02324-0.

[25] M. Afkanpour, E. Hosseinzadeh, and H. Tabesh, "Identify the most appropriate imputation method for handling missing values in clinical structured datasets: a systematic review," *BMC Med. Res. Methodol.*, vol. 24, no. 1, p. 188, Aug. 2024, doi: 10.1186/s12874-024-02310-6.