STUDI KOMPARATIF ALGORITMA RANDOM FOREST, GRADIENT BOOSTING, DAN LIGHTGBM PADA KLASIFIKASI RISIKO DIABETES MELITUS

Mohamad Rizal Maulana, Setyawan Wibisono
Teknik Informatika, Universitas Stikubank Semarang
Jl. Kendeng V Bendan Ngisor, Semarang 50233, Indonesia
mohamadrizal0024@mhs.unisbank.ac.id

ABSTRAK

Diabetes melitus merupakan penyakit kronis yang menunjukkan kecenderungan peningkatan prevalensi secara berkelanjutan dan telah menjadi permasalahan kesehatan yang signifikan, baik pada tingkat nasional maupun global. Ketidakmampuan dalam mendeteksi penyakit ini sejak dini dapat meningkatkan risiko munculnya berbagai komplikasi serius. Salah satu faktor yang berkontribusi terhadap keterlambatan deteksi dini diabetes melitus adalah masih dominannya penggunaan metode diagnosis berbasis pemeriksaan laboratorium, yang memerlukan biaya relatif tinggi serta waktu pemeriksaan yang cukup lama, sehingga belum sepenuhnya menjangkau seluruh lapisan masyarakat.Penelitian ini bertujuan untuk mengevaluasi dan membandingkan kinerja algoritma Random Forest, Gradient Boosting, dan Light Gradient Boosting Machine (LightGBM) dalam mengklasifikasikan risiko diabetes melitus. Pendekatan penelitian yang digunakan bersifat eksperimental dengan memanfaatkan dataset diabetes yang bersumber dari UCI Machine Learning Repository. Proses penelitian meliputi tahap prapemrosesan data, pembangunan model klasifikasi menggunakan algoritma *ensemble learning*, serta pengujian performa model melalui metode 10-Fold Cross Validation. Kinerja model dievaluasi menggunakan metrik akurasi, presisi, recall, F1-score, dan *Area Under Curve* (AUC).Hasil analisis menunjukkan bahwa algoritma Gradient Boosting memberikan performa paling baik dibandingkan Random Forest dan LightGBM berdasarkan nilai evaluasi yang diperoleh. Temuan ini menunjukkan bahwa Gradient Boosting memiliki potensi yang signifikan untuk dikembangkan sebagai dasar sistem deteksi dini diabetes melitus berbasis *machine learning*.

Kata kunci : *diabetes melitus, machine learning, klasifikasi, ensemble learning, gradient boosting, deteksi dini*

1. PENDAHULUAN

Diabetes melitus merupakan salah satu penyakit tidak menular yang masih menjadi tantangan utama dalam sistem kesehatan di Indonesia. Prevalensi diabetes menunjukkan tren peningkatan yang signifikan, seiring dengan perubahan gaya hidup masyarakat, seperti pola konsumsi makanan tidak sehat dan rendahnya tingkat aktivitas fisik [1]. Peningkatan jumlah penderita diabetes tersebut berdampak langsung terhadap meningkatnya angka kesakitan serta beban pembiayaan layanan kesehatan nasional [2].

Permasalahan diabetes melitus tidak hanya berkaitan dengan tingginya prevalensi, tetapi juga risiko komplikasi yang menyertainya. Diabetes yang tidak terdeteksi secara dini dan tidak dikelola dengan baik berpotensi menimbulkan berbagai komplikasi kronis, antara lain penyakit kardiovaskular, gangguan ginjal, neuropati, gangguan penglihatan, hingga amputasi [3] [4]. Beberapa penelitian melaporkan bahwa sebagian besar penderita baru mengetahui kondisi diabetesnya setelah komplikasi berkembang, yang mengindikasikan masih rendahnya efektivitas upaya deteksi dini di masyarakat [5]. Di sisi lain, metode diagnosis diabetes yang umum digunakan saat ini masih bergantung pada pemeriksaan laboratorium, yang memerlukan biaya dan waktu serta belum sepenuhnya mudah diakses oleh seluruh lapisan masyarakat [6].

Perkembangan teknologi kecerdasan buatan, khususnya machine learning, memberikan peluang dalam mendukung upaya deteksi dini diabetes melitus. Pendekatan machine learning memungkinkan pemanfaatan data kesehatan individu untuk melakukan prediksi dan klasifikasi penyakit secara cepat dan efisien [7]. Sejumlah penelitian di Indonesia menunjukkan bahwa penerapan algoritma machine learning pada klasifikasi diabetes mampu menghasilkan tingkat akurasi yang baik dan berpotensi digunakan sebagai sistem pendukung pengambilan keputusan medis [8] [9].

Algoritma berbasis ensemble learning, seperti Random Forest, Gradient Boosting, dan Light Gradient Boosting Machine (LightGBM), banyak diterapkan dalam klasifikasi data kesehatan karena kemampuannya dalam menangani data kompleks dan meningkatkan akurasi prediksi [10]. Random Forest bekerja dengan mengombinasikan sejumlah pohon keputusan yang dibangun secara paralel, sedangkan Gradient Boosting membangun model secara bertahap dengan menekankan perbaikan kesalahan prediksi pada tahap sebelumnya [11]. LightGBM merupakan pengembangan dari metode Gradient Boosting yang dirancang untuk meningkatkan efisiensi komputasi dan mempercepat proses pelatihan model [12]. Perbedaan mekanisme kerja tersebut menyebabkan variasi kinerja algoritma ketika diterapkan pada dataset kesehatan yang berbeda [13].

Meskipun ketiga algoritma tersebut telah banyak digunakan dalam penelitian sebelumnya, kajian yang membandingkan kinerja Random Forest, Gradient Boosting, dan LightGBM secara langsung pada kasus klasifikasi diabetes dengan menggunakan dataset serta metode evaluasi yang sama masih relatif terbatas [14]. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja ketiga algoritma ensemble learning tersebut dalam klasifikasi diabetes melitus.

Penelitian ini dilakukan melalui tahapan pengumpulan dataset diabetes, prapemrosesan data, penerapan algoritma Random Forest, Gradient Boosting, dan LightGBM, serta evaluasi performa model menggunakan metrik akurasi, precision, recall, F1-score, dan area under curve (AUC). Hasil penelitian diharapkan dapat memberikan rekomendasi algoritma yang paling optimal sebagai dasar pengembangan sistem deteksi dini diabetes melitus [15].

2. TINJAUAN PUSTAKA

Dalam beberapa tahun terakhir, pemanfaatan *machine learning* untuk klasifikasi dan prediksi diabetes di Indonesia menunjukkan perkembangan yang signifikan. Salah satu penelitian membandingkan kinerja berbagai algoritma, termasuk Random Forest dan Gradient Boosting, dalam memprediksi diabetes, dan melaporkan bahwa Gradient Boosting menghasilkan nilai ROC AUC yang lebih tinggi dibandingkan Random Forest maupun Support Vector Classifier [1]. Hasil tersebut mengindikasikan bahwa pendekatan *ensemble learning* memiliki kemampuan yang baik dalam memodelkan kompleksitas pola data medis.

Studi lain yang mengkaji penerapan Random Forest pada dataset kesehatan lokal menekankan pentingnya pengujian kinerja model klasifikasi dalam konteks penggunaan nyata. Susanto *et al*. melaporkan bahwa algoritma Random Forest mampu mencapai tingkat akurasi dan metrik evaluasi lain yang tinggi dalam prediksi diabetes pada data uji tertentu, sekaligus memberikan implikasi praktis bagi implementasi di lingkungan klinis [2]. Temuan tersebut turut menegaskan peran *machine learning* dalam mengatasi keterbatasan metode diagnosis konvensional yang cenderung membutuhkan waktu dan biaya besar.

Penelitian lokal lainnya turut mengeksplorasi penerapan teknik prapemrosesan data, seperti imputasi, serta penggunaan algoritma LightGBM dalam konteks prediksi diabetes. Pendekatan tersebut terbukti efektif dalam menangani data dengan nilai hilang dan menghasilkan kinerja prediksi yang kompetitif, ditunjukkan oleh nilai akurasi dan AUC yang memadai [3]. Temuan ini memperkuat pemahaman bahwa algoritma *gradient boosting* generasi baru seperti LightGBM berpotensi menjadi alternatif yang andal dalam pengembangan sistem deteksi dini berbasis data.

Studi lain melakukan perbandingan langsung antara Random Forest dan LightGBM dalam klasifikasi diabetes melitus gestasional. Hasil penelitian menunjukkan bahwa kedua algoritma mampu melakukan klasifikasi risiko secara efektif, namun karakteristik dataset dan pengaturan parameter model berperan signifikan terhadap kinerja akhir yang dihasilkan [4]. Temuan ini menegaskan pentingnya pemilihan serta penyesuaian algoritma sesuai dengan konteks data kesehatan yang dianalisis.

3. METODE PENELITIAN

3.1. Objek Penelitian

Penelitian ini menggunakan dataset *Diabetes Health Indicators* yang bersumber dari UCI Machine Learning Repository sebagai objek kajian. Dataset tersebut dimanfaatkan sebagai data utama dalam pengembangan dan evaluasi model klasifikasi untuk memprediksi status diabetes. Variabel *Diabetes_012* terdiri atas tiga kelas, yaitu 0 untuk individu tanpa diabetes, 1 untuk kondisi pra-diabetes, dan 2 untuk penderita diabetes. Pemilihan dataset ini didasarkan pada ketersediaan jumlah data yang besar, keberagaman fitur yang relevan, serta karakteristik data yang sesuai untuk penerapan algoritma klasifikasi berbasis *ensemble learning*.

3.2. Ensemble Learning

Ensemble learning dalam *machine learning* merupakan pendekatan pemodelan yang mengombinasikan beberapa model pembelajaran untuk memperoleh hasil prediksi yang lebih andal. Prinsip dasarnya adalah memanfaatkan keberagaman karakteristik antar model sehingga kesalahan pada satu model dapat diminimalkan melalui kontribusi model lainnya. Pendekatan ini banyak digunakan pada tugas klasifikasi karena mampu meningkatkan akurasi, menekan risiko *overfitting*, dan menghasilkan kinerja yang lebih stabil pada data baru.

Penelitian ini menerapkan tiga algoritma *ensemble learning*, yaitu Random Forest, Gradient Boosting, dan Light Gradient Boosting Machine (LightGBM). Ketiga metode tersebut dipilih karena kemampuannya dalam mengelola data dengan jumlah fitur yang relatif besar serta memodelkan pola hubungan nonlinier yang umum dijumpai pada data kesehatan, khususnya dalam tugas klasifikasi diabetes.

Random Forest merupakan algoritma *ensemble learning* yang membangun sejumlah pohon keputusan secara independen melalui proses pengambilan sampel acak pada data dan fitur. Setiap pohon menghasilkan prediksi tersendiri, yang kemudian digabungkan menggunakan mekanisme *majority voting* untuk menentukan hasil akhir. Strategi ini membuat Random Forest lebih tahan terhadap *noise* dan *overfitting*, serta mampu memberikan kinerja yang stabil pada data dengan tingkat kompleksitas tinggi.

Berbeda dari Random Forest, Gradient Boosting menyusun model secara iteratif, di mana setiap model selanjutnya diarahkan untuk memperbaiki kesalahan

prediksi pada tahap sebelumnya. Proses pelatihan dilakukan melalui optimasi fungsi kerugian berbasis gradien, sehingga akurasi model meningkat secara bertahap. Pendekatan ini memungkinkan Gradient Boosting menghasilkan kinerja prediksi yang tinggi dan efektif dalam menangkap pola data yang kompleks, khususnya pada studi prediksi penyakit.

Light Gradient Boosting Machine (LightGBM) merupakan pengembangan dari algoritma Gradient Boosting yang difokuskan pada peningkatan efisiensi komputasi dan percepatan proses pelatihan. Algoritma ini menggunakan pendekatan pembelajaran berbasis histogram serta strategi pertumbuhan pohon *leaf-wise*, sehingga mampu mengurangi kebutuhan sumber daya sekaligus mempercepat pelatihan model. Meskipun demikian, LightGBM tetap mempertahankan tingkat akurasi yang tinggi, sehingga sesuai diterapkan pada dataset berskala besar dan permasalahan klasifikasi, termasuk prediksi diabetes.

3.3. Metode Pengembangan

Penelitian ini menerapkan pendekatan *eksperimental* komputasional dalam pengembangan dan evaluasi model klasifikasi diabetes. Pendekatan tersebut digunakan untuk menguji kinerja algoritma *machine learning* serta memperoleh model dengan kemampuan prediktif yang optimal. Seluruh tahapan penelitian dilaksanakan secara terstruktur dan disajikan pada Gambar 1.

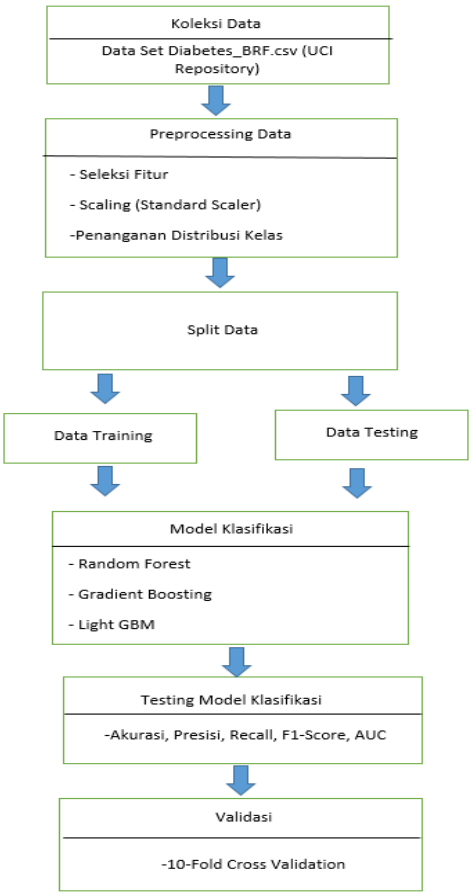
Penelitian ini diawali dengan studi literatur yang membahas diabetes melitus, metode klasifikasi, serta pendekatan *ensemble learning* guna memahami urgensi deteksi dini dalam mencegah komplikasi penyakit.

Tahap selanjutnya adalah pengumpulan data menggunakan dataset publik dari UCI Machine Learning Repository yang memuat informasi kesehatan dan karakteristik demografis terkait diabetes. Penggunaan dataset publik bertujuan untuk menjamin transparansi data serta memudahkan perbandingan hasil dengan penelitian sebelumnya.

Setelah itu, dilakukan prapemrosesan data untuk meningkatkan kualitas dataset, meliputi pembersihan data, penanganan nilai hilang atau tidak valid, serta transformasi data seperti standarisasi fitur numerik agar berada pada skala yang seragam. Tahap ini bertujuan memastikan kesiapan data sebelum pemodelan.

Pada fase pengembangan model, diterapkan algoritma Random Forest, Gradient Boosting, dan Light Gradient Boosting Machine (LightGBM) untuk membangun model klasifikasi status diabetes. Keandalan model diuji menggunakan metode 10-Fold Cross Validation. Selanjutnya, kinerja model dievaluasi berdasarkan metrik akurasi, presisi, recall, F1-score, dan Area Under Curve (AUC) guna menilai kemampuan klasifikasi secara menyeluruh.

Tahap akhir mencakup analisis hasil dan penarikan kesimpulan untuk menentukan model dengan kinerja terbaik serta memberikan rekomendasi bagi penelitian lanjutan dalam mendukung pengembangan sistem deteksi dini diabetes melitus.



Gambar 1. Tahapan penelitian

Gambar tersebut menggambarkan tahapan metodologi penelitian klasifikasi diabetes berbasis *machine learning*. Proses dimulai dengan pengumpulan data menggunakan dataset *Diabetes_BRFSS.csv* yang diperoleh dari UCI Repository. Tahap prapemrosesan mencakup seleksi fitur, normalisasi data menggunakan *Standard Scaler*, serta penanganan ketidakseimbangan kelas.

Selanjutnya, data dibagi menjadi data pelatihan dan pengujian untuk keperluan pembentukan dan evaluasi model. Tiga algoritma *ensemble learning*—Random Forest, Gradient Boosting, dan LightGBM—diterapkan pada tahap klasifikasi. Kinerja model dinilai menggunakan metrik akurasi, presisi, recall, F1-score, dan Area Under Curve (AUC), kemudian divalidasi melalui metode 10-Fold Cross Validation guna memastikan stabilitas dan kemampuan generalisasi model.

4. HASIL DAN PEMBAHASAN

4.1. Evaluasi Model Klasifikasi

Hasil eksperimen menyajikan evaluasi kinerja tiga model klasifikasi, yaitu Random Forest, Gradient Boosting, dan LightGBM, menggunakan metrik akurasi, presisi, recall, F1-score, serta Area Under

Curve (AUC). Ringkasan performa masing-masing algoritma ditampilkan pada Tabel 1.

Tabel 1. Hasil Eksperiman

Model	Akurasi	Presisi	Recall	F1-Score	AUC
Random Forest	0.847839	0.802001	0.847839	0.797144	0.817302
Gradient Boosting	0.849120	0.804406	0.849120	0.810172	0.822206
LightGBM	0.849337	0.805089	0.849337	0.811304	0.822696

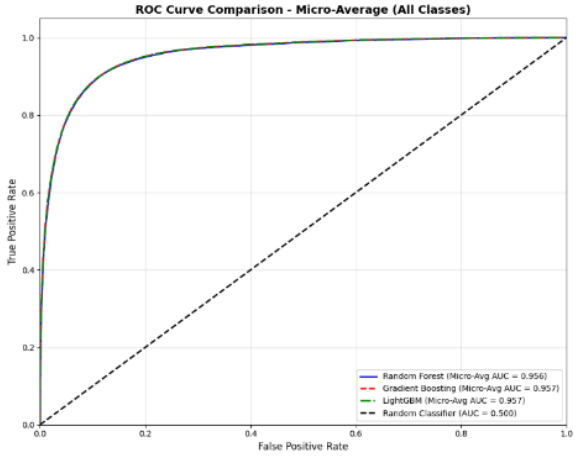
Model Random Forest menunjukkan kinerja yang relatif baik dalam klasifikasi diabetes dengan akurasi sebesar 0,84783979 (84,78%). Nilai presisi sebesar 0,80200183 menunjukkan bahwa sebagian besar prediksi positif bersifat tepat, sedangkan recall sebesar 0,84783979 mencerminkan kemampuan model dalam mengidentifikasi data positif secara memadai. F1-score sebesar 0,79714429 mengindikasikan keseimbangan yang cukup antara presisi dan recall, sementara nilai AUC sebesar 0,81730225 menunjukkan kemampuan diskriminatif yang baik. Secara umum, Random Forest menghasilkan performa yang stabil dan dapat diandalkan.

Gradient Boosting memperlihatkan peningkatan kinerja pada hampir seluruh metrik evaluasi. Model ini mencapai akurasi 0,84912093 (84,91%), dengan presisi sebesar 0,80440644 dan recall sebesar 0,84912093, yang menunjukkan efektivitas lebih tinggi dalam mengurangi *false positive* serta mendeteksi kelas positif. Nilai F1-score sebesar 0,81017229 dan AUC sebesar 0,82220648 mengindikasikan keseimbangan klasifikasi serta kemampuan pemisahan kelas yang lebih baik dibandingkan Random Forest.

LightGBM memberikan performa terbaik di antara ketiga model, meskipun selisih peningkatannya relatif kecil. Algoritma ini memperoleh akurasi tertinggi sebesar 0,84933774 (84,93%), dengan presisi 0,80508932 dan recall 0,84933774. Nilai F1-score sebesar 0,81130409 menunjukkan keseimbangan presisi dan recall yang paling optimal, sedangkan AUC tertinggi sebesar 0,82269624 menegaskan kemampuan LightGBM dalam membedakan kelas positif dan negatif. Secara keseluruhan, LightGBM merupakan model dengan kinerja paling optimal pada tugas klasifikasi diabetes dalam penelitian ini.

4.2. ROC Receiver Operating Characteristic)

Gambar 2 menyajikan kurva ROC (*Receiver Operating Characteristic*) yang membandingkan kinerja algoritma Random Forest, Gradient Boosting, dan LightGBM menggunakan pendekatan *micro-average* untuk seluruh kelas. Kurva ROC digunakan untuk menilai kemampuan masing-masing model dalam membedakan kelas positif dan negatif pada berbagai nilai ambang keputusan.



Gambar 2. ROC

Pada kurva ROC, sumbu horizontal merepresentasikan *False Positive Rate* (FPR), yaitu proporsi data negatif yang keliru diklasifikasikan sebagai positif, sedangkan sumbu vertikal menunjukkan *True Positive Rate* (TPR) atau sensitivitas, yang menggambarkan tingkat keberhasilan model dalam mengidentifikasi data positif. Nilai FPR dan TPR berada pada rentang 0 hingga 1. Garis diagonal putus-putus pada grafik menggambarkan kinerja klasifikator acak dengan nilai AUC sebesar 0,5, di mana kurva yang mendekati garis tersebut menunjukkan kemampuan diskriminasi yang rendah. Sebaliknya, kurva yang semakin mendekati sudut kiri atas menandakan performa klasifikasi yang semakin baik.

Berdasarkan hasil visualisasi, ketiga model menghasilkan kurva ROC yang berada jauh di atas garis acuan, menunjukkan kemampuan klasifikasi yang sangat baik. Nilai AUC *micro-average* masing-masing model adalah 0,956 untuk Random Forest, serta 0,957 untuk Gradient Boosting dan LightGBM, yang mengindikasikan tingkat diskriminasi yang sangat tinggi.

Perbedaan nilai AUC antar model relatif kecil dan kurva ROC tampak hampir saling tumpang tindih, menandakan bahwa kemampuan diskriminasi ketiga algoritma pada berbagai nilai ambang bersifat sebanding. Secara keseluruhan, hasil ini menunjukkan bahwa Random Forest, Gradient Boosting, dan LightGBM merupakan metode yang efektif dan andal dalam klasifikasi diabetes pada dataset yang digunakan, dengan kinerja pemisahan kelas yang hampir setara.

4.3. Cross-Validation

Cross-validation merupakan metode evaluasi untuk menilai kinerja model *machine learning* dengan membagi dataset ke dalam beberapa bagian (*fold*). Pada skema 10-Fold Cross Validation, data dipartisi menjadi 10 subset, di mana pada setiap iterasi model dilatih menggunakan sembilan subset dan diuji pada satu subset yang berbeda. Proses ini diulang hingga seluruh subset digunakan sebagai data uji, sehingga

mampu menekan risiko *overfitting* dan menghasilkan estimasi kinerja model yang lebih stabil serta representatif.

Tabel 2. *10-Fold Cross Validation*

Model	Akurasi	Presisi	Recall	F1-Score	AUC
Random Forest	0.84784	0.802002	0.84784	0.797144	0.817302
Gradient Boosting	0.849338	0.805089	0.849338	0.811304	0.822696
LightGBM	0.849121	0.804406	0.849121	0.810172	0.822206

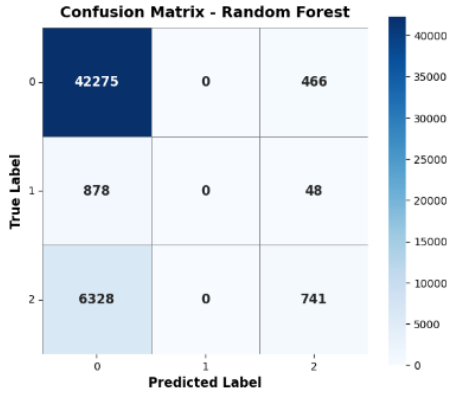
Berdasarkan hasil 10-Fold Cross Validation, Random Forest memperoleh akurasi sebesar 84,78% dengan presisi 80,20% dan recall 84,78%, yang menunjukkan kemampuan klasifikasi yang cukup baik. Nilai F1-score sebesar 79,71% mencerminkan keseimbangan presisi dan recall, sementara AUC sebesar 81,73% mengindikasikan kemampuan pemisahan kelas yang memadai.

Gradient Boosting menunjukkan kinerja paling unggul dengan akurasi tertinggi sebesar 84,93%, presisi 80,51%, dan recall 84,93%. Keseimbangan klasifikasi tercermin dari F1-score sebesar 81,13%, sedangkan nilai AUC tertinggi sebesar 82,27% menegaskan kemampuan diskriminasi kelas yang optimal.

LightGBM menampilkan performa yang kompetitif dengan akurasi 84,91%, presisi 80,44%, dan recall 84,91%. Nilai F1-score sebesar 81,02% serta AUC sebesar 82,22% menunjukkan bahwa kinerjanya hampir setara dengan Gradient Boosting, meskipun sedikit lebih rendah pada beberapa metrik.

Secara keseluruhan, Gradient Boosting memberikan performa terbaik berdasarkan akurasi, F1-score, dan AUC. LightGBM menjadi alternatif yang kuat dengan kinerja yang mendekati, sementara Random Forest tetap menunjukkan stabilitas dan keandalan meskipun berada sedikit di bawah kedua model lainnya.

4.4. Confussion Matrix Random Forest



Gambar 3. *Confussion Matrix* Random Forest

Confusion Matrix merupakan alat evaluasi yang digunakan untuk memahami performa model klasifikasi dengan menampilkan jumlah prediksi yang benar dan salah pada setiap kelas. Gambar 3 menunjukkan confusion matrix untuk algoritma Random Forest pada klasifikasi risiko Diabetes Melitus dengan tiga kelas, yaitu kelas 0, kelas 1, dan kelas 2.

Confusion matrix berbentuk persegi dengan tiga baris dan tiga kolom yang masing-masing merepresentasikan label aktual dan label prediksi. Sumbu vertikal (True Label) menunjukkan kelas sebenarnya dari data, sedangkan sumbu horizontal (Predicted Label) menunjukkan kelas hasil prediksi model. Setiap sel pada matriks menggambarkan jumlah data yang diklasifikasikan berdasarkan kombinasi antara label aktual dan label prediksi.

Berikut adalah penjelasan dari masing-masing elemen pada confusion matrix tersebut:

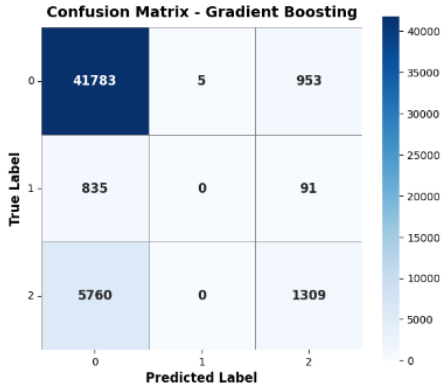
a. Kelas 0 (Risiko Rendah)
   Sebanyak 42.275 data dengan label asli kelas 0 berhasil diprediksi dengan benar sebagai kelas 0 oleh model. Namun, terdapat 466 data dari kelas 0 yang salah diklasifikasikan sebagai kelas 2, dan tidak ada data yang diprediksi sebagai kelas 1. Hal ini menunjukkan bahwa model Random Forest sangat baik dalam mengenali kelas 0.
b. Kelas 1 (Risiko Sedang)
   Untuk kelas 1, hanya 48 data yang berhasil diklasifikasikan dengan benar. Sebaliknya, terdapat 878 data kelas 1 yang salah diprediksi sebagai kelas 0, dan tidak ada data yang diprediksi sebagai kelas 1 oleh model. Kondisi ini menunjukkan bahwa model mengalami kesulitan dalam membedakan kelas 1 dari kelas lainnya.
c. Kelas 2 (Risiko Tinggi)
   Pada kelas 2, model berhasil mengklasifikasikan 741 data dengan benar. Namun, masih terdapat 6.328 data kelas 2 yang salah diklasifikasikan sebagai kelas 0, serta tidak ada prediksi ke kelas 1. Hal ini mengindikasikan bahwa sebagian besar kesalahan klasifikasi pada kelas 2 cenderung diprediksi sebagai kelas 0.

Berdasarkan hasil confusion matrix tersebut, dapat disimpulkan bahwa model Random Forest memiliki performa yang sangat baik dalam mengklasifikasikan kelas 0, namun kurang optimal dalam mengenali kelas 1 dan kelas 2. Model cenderung memprediksi data ke kelas mayoritas (kelas 0), yang mengindikasikan adanya ketidakseimbangan kelas *(class imbalance)* pada dataset. Kondisi ini berpotensi meningkatkan risiko kesalahan klasifikasi, terutama pada kelas risiko sedang dan tinggi, yang secara klinis lebih penting. Oleh karena itu, diperlukan strategi lanjutan seperti penyeimbangan data atau penyesuaian parameter model untuk meningkatkan kemampuan klasifikasi pada kelas minoritas.

4.5. Confussion Matrix Gradient Boosting

Gambar 4 menampilkan confusion matrix hasil evaluasi model Gradient Boosting dengan tiga kelas (0, 1, dan 2). Matriks ini menunjukkan perbandingan antara label sebenarnya *(True Label)* dan label hasil

prediksi model *(Predicted Label)*, sehingga dapat digunakan untuk menilai performa klasifikasi model secara lebih rinci.



Gambar 4. *Confussion Matrix* Gradient Boosting

Berikut adalah penjelasan dari masing-masing elemen pada confusion matrix tersebut:

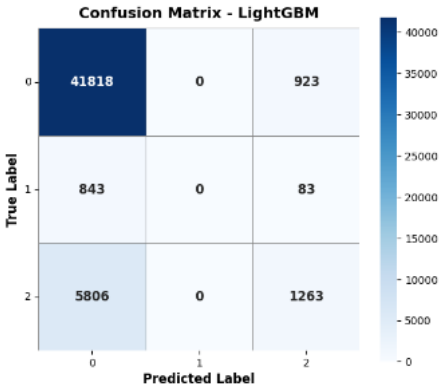
a. Kelas 0 (Risiko Rendah)
   Sebanyak 41.783 data dengan label asli kelas 0 berhasil diprediksi dengan benar oleh model sebagai kelas 0. Hal ini menunjukkan bahwa model Gradient Boosting memiliki kemampuan yang sangat baik dalam mengenali data pada kelas risiko rendah.
b. Kelas 1 (Risiko Sedang)
   Pada kelas 1, performa model terlihat kurang optimal. Tidak ada satu pun data yang berhasil diprediksi dengan benar sebagai kelas 1. Sebanyak 835 data justru diklasifikasikan sebagai kelas 0, sementara 91 data lainnya diprediksi sebagai kelas 2.
c. Kelas 2 (Risiko Tinggi)
   Untuk kelas 2, model berhasil mengklasifikasikan 1.309 data dengan benar. Namun, jumlah kesalahan masih cukup besar, terutama karena 5.760 data kelas 2 salah diprediksi sebagai kelas 0. Tidak terdapat data yang diprediksi sebagai kelas 1.

Hal ini menunjukkan bahwa model cenderung mengelompokkan data kelas 2 ke dalam kelas 0, yang mengindikasikan adanya kecenderungan bias terhadap kelas mayoritas.

Secara keseluruhan, hasil confusion matrix menunjukkan bahwa model Gradient Boosting memiliki performa yang sangat baik dalam mengenali kelas 0, tetapi kurang optimal dalam mendeteksi kelas 1 dan kelas 2. Model tampak memiliki kecenderungan bias terhadap kelas mayoritas, yang menyebabkan tingginya jumlah kesalahan pada kelas minoritas. Oleh karena itu, diperlukan upaya perbaikan seperti penyeimbangan data atau penyesuaian parameter model agar kemampuan klasifikasi menjadi lebih merata dan akurat.

4.6. Confussion Matrix LightGBM

Dalam penelitian ini, confusion matrix digunakan untuk mengevaluasi kinerja model LightGBM dalam mengklasifikasikan tingkat risiko, sebagaimana ditunjukkan pada Gambar 5.



Gambar 5. *Confussion Matrix* LightGBM

Confusion matrix ini menggambarkan perbandingan antara label sebenarnya dan hasil prediksi model untuk tiga kelas, yaitu kelas 0 (risiko rendah), kelas 1 (risiko sedang), dan kelas 2 (risiko tinggi). Melalui matriks ini, dapat dianalisis sejauh mana model mampu melakukan klasifikasi secara akurat serta mengidentifikasi jenis kesalahan yang terjadi. Dalam penelitian ini, confusion matrix digunakan untuk mengevaluasi kinerja model LightGBM dalam mengklasifikasikan tingkat risiko, sebagaimana ditunjukkan pada Gambar 5. Confusion matrix ini menggambarkan perbandingan antara label sebenarnya dan hasil prediksi model untuk tiga kelas, yaitu kelas 0 (risiko rendah), kelas 1 (risiko sedang), dan kelas 2 (risiko tinggi). Melalui matriks ini, dapat dianalisis sejauh mana model mampu melakukan klasifikasi secara akurat serta mengidentifikasi jenis kesalahan yang terjadi.

Berikut adalah penjelasan masing-masing elemen pada confusion matrix tersebut:

a. Kelas 0 (Risiko Rendah)
   Pada kelas 0, model LightGBM berhasil mengklasifikasikan 41.818 data dengan benar sebagai kelas 0. Hasil ini menunjukkan bahwa model memiliki performa yang sangat baik dalam mengenali data dengan risiko rendah.
b. Kelas 1 (Risiko Sedang
   Pada kelas 1, performa model tergolong kurang optimal. Tidak terdapat satu pun data yang berhasil diklasifikasikan dengan benar sebagai kelas 1. Sebanyak 843 data justru diprediksi sebagai kelas 0, sementara 83 data lainnya diklasifikasikan sebagai kelas 2.
c. Kelas 2 (Risiko Tinggi)
   Untuk kelas 2, model berhasil mengklasifikasikan 1.263 data dengan benar. Namun demikian, masih terdapat 5.806 data yang salah diklasifikasikan sebagai kelas 0, dan tidak ada data yang diprediksi sebagai kelas 1.

Berdasarkan *confusion matrix* pada Gambar 5, kinerja Random Forest, Gradient Boosting, dan LightGBM dalam mengklasifikasikan tiga kelas risiko dapat dianalisis secara komparatif. Random Forest menunjukkan performa sangat baik pada kelas 0 (risiko rendah), dengan 42.275 data terklasifikasi benar dan tingkat kesalahan yang rendah, yaitu 466 data salah diprediksi sebagai kelas 2 tanpa kesalahan ke kelas 1. Namun, model ini kurang efektif dalam mengenali kelas minoritas. Pada kelas 1, hanya 48 data yang terklasifikasi benar, sementara 878 data salah diprediksi sebagai kelas 0. Demikian pula pada kelas 2, meskipun terdapat 741 prediksi benar, sebanyak 6.328 data masih salah diklasifikasikan ke kelas mayoritas, menunjukkan kecenderungan bias terhadap kelas dominan.

Gradient Boosting memperlihatkan pola yang serupa, dengan 41.783 data kelas 0 terklasifikasi secara tepat dan kesalahan yang relatif kecil pada kelas lain. Akan tetapi, model ini gagal mengidentifikasi kelas 1 secara benar, dengan 835 data salah diprediksi sebagai kelas 0 dan 91 data sebagai kelas 2. Pada kelas 2, sebanyak 1.309 data berhasil dikenali, meskipun 5.760 data masih keliru diklasifikasikan sebagai kelas 0, mengindikasikan keterbatasan dalam mendeteksi kelas minoritas.

LightGBM juga menunjukkan performa yang sebanding, dengan 41.818 data kelas 0 diklasifikasikan secara tepat. Namun, tidak terdapat prediksi benar pada kelas 1, dengan 843 data salah dipetakan ke kelas 0 dan 83 data ke kelas 2. Pada kelas 2, LightGBM menghasilkan 1.263 prediksi benar, tetapi masih terdapat 5.806 kesalahan ke kelas mayoritas. Secara umum, ketiga model menunjukkan kecenderungan bias terhadap kelas dominan, meskipun LightGBM menawarkan keunggulan dalam efisiensi dan kecepatan komputasi.

5. KESIMPULAN DAN SARAN

Penelitian ini bertujuan menganalisis kinerja algoritma Random Forest, Gradient Boosting, dan LightGBM dalam klasifikasi status diabetes menggunakan dataset *diabetes_BRF* melalui tahapan pembagian data, pelatihan, pengujian, dan *cross-validation*. Hasil penelitian menunjukkan bahwa pendekatan *machine learning* mampu menghasilkan performa klasifikasi yang baik meskipun dataset memiliki distribusi kelas yang tidak seimbang. Ketiga algoritma dapat mempelajari pola data secara efektif, namun Random Forest masih memiliki keterbatasan dalam membedakan kelas tertentu, sedangkan Gradient Boosting menunjukkan peningkatan kinerja yang belum sepenuhnya konsisten. LightGBM memberikan hasil paling optimal dengan nilai akurasi, F1-score, dan AUC yang lebih stabil serta kemampuan generalisasi yang baik, sehingga dinilai paling sesuai untuk tugas klasifikasi diabetes pada studi ini.

Meskipun demikian, penelitian ini memiliki sejumlah keterbatasan, antara lain penggunaan parameter model bawaan, belum diterapkannya teknik penyeimbangan kelas, keterbatasan analisis interpretabilitas, serta belum adanya implementasi dalam bentuk aplikasi. Oleh karena itu, penelitian selanjutnya disarankan untuk melakukan optimasi parameter, penanganan ketidakseimbangan data, penerapan metode *explainable AI*, pengembangan sistem pendukung keputusan, serta pengujian pada dataset dan skema validasi yang lebih beragam.

DAFTAR PUSTAKA

[1] P. Soewondo dan A. Tahapary, “Tantangan dan strategi penanggulangan diabetes melitus di Indonesia,” *Jurnal Penyakit Dalam Indonesia*, vol. 8, no. 2, pp. 65–72, 2021.

[2] S. Trisnawati dan S. Setyorogo, “Faktor risiko kejadian diabetes melitus tipe 2 di Indonesia,” *Jurnal Kesehatan Masyarakat*, vol. 16, no. 1, pp. 1–9, 2021.

[3] S. Waspadji, “Komplikasi kronik diabetes melitus,” *Majalah Kedokteran Indonesia*, vol. 71, no. 2, pp. 85–92, 2021.

[4] A. Decroli, “Komplikasi dan pencegahan diabetes melitus tipe 2,” *Jurnal Kesehatan Andalas*, vol. 10, no. 1, pp. 1–8, 2021.

[5] R. Hidayat dan Y. Nugroho, “Deteksi dini diabetes melitus berbasis data kesehatan,” *Jurnal Teknologi dan Sistem Komputer*, vol. 10, no. 2, pp. 78–85, 2022.

[6] Suyono, “Diagnosis dan penatalaksanaan diabetes melitus,” *Majalah Kedokteran Indonesia*, vol. 71, no. 1, pp. 45–52, 2021.

[7] E. Prasetyo, “Penerapan machine learning dalam bidang kesehatan,” *Jurnal Teknologi Informasi*, vol. 7, no. 2, pp. 101–109, 2021.

[8] A. D. Putra dan N. Sari, “Prediksi diabetes menggunakan algoritma machine learning,” *Jurnal Ilmiah Teknik Informatika*, vol. 15, no. 2, pp. 85–92, 2021.

[9] F. Rahman, R. Hidayat, dan Y. Nugroho, “Klasifikasi risiko diabetes menggunakan machine learning,” *Jurnal RESTI*, vol. 6, no. 1, pp. 45–52, 2022.

[10] B. Santoso dan A. Nugroho, “Implementasi Random Forest untuk klasifikasi penyakit,” *Jurnal RESTI*, vol. 5, no. 4, pp. 712–719, 2021.

[11] Y. Nugroho, “Penerapan Gradient Boosting pada klasifikasi data medis,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 3, pp. 523–530, 2021.

[12] R. A. Wijaya, “Implementasi LightGBM pada klasifikasi data kesehatan,” *Jurnal Teknologi Informasi dan Komunikasi*, vol. 11, no. 2, pp. 134–142, 2022.

[13] R. Hidayat dan Y. Nugroho, “Perbandingan algoritma machine learning pada data kesehatan,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 1, pp. 15–23, 2022.

[14] R. M. Putri, “Perbandingan algoritma ensemble learning pada data kesehatan,” *Jurnal Informatika Mulawarman*, vol. 16, no. 2, pp. 101–108, 2021.

[15] S. Nugraha et al., “Evaluasi performa algoritma machine learning pada data medis,” *Prosiding Seminar Nasional Teknologi Informasi*, pp. 210–217, 2023.