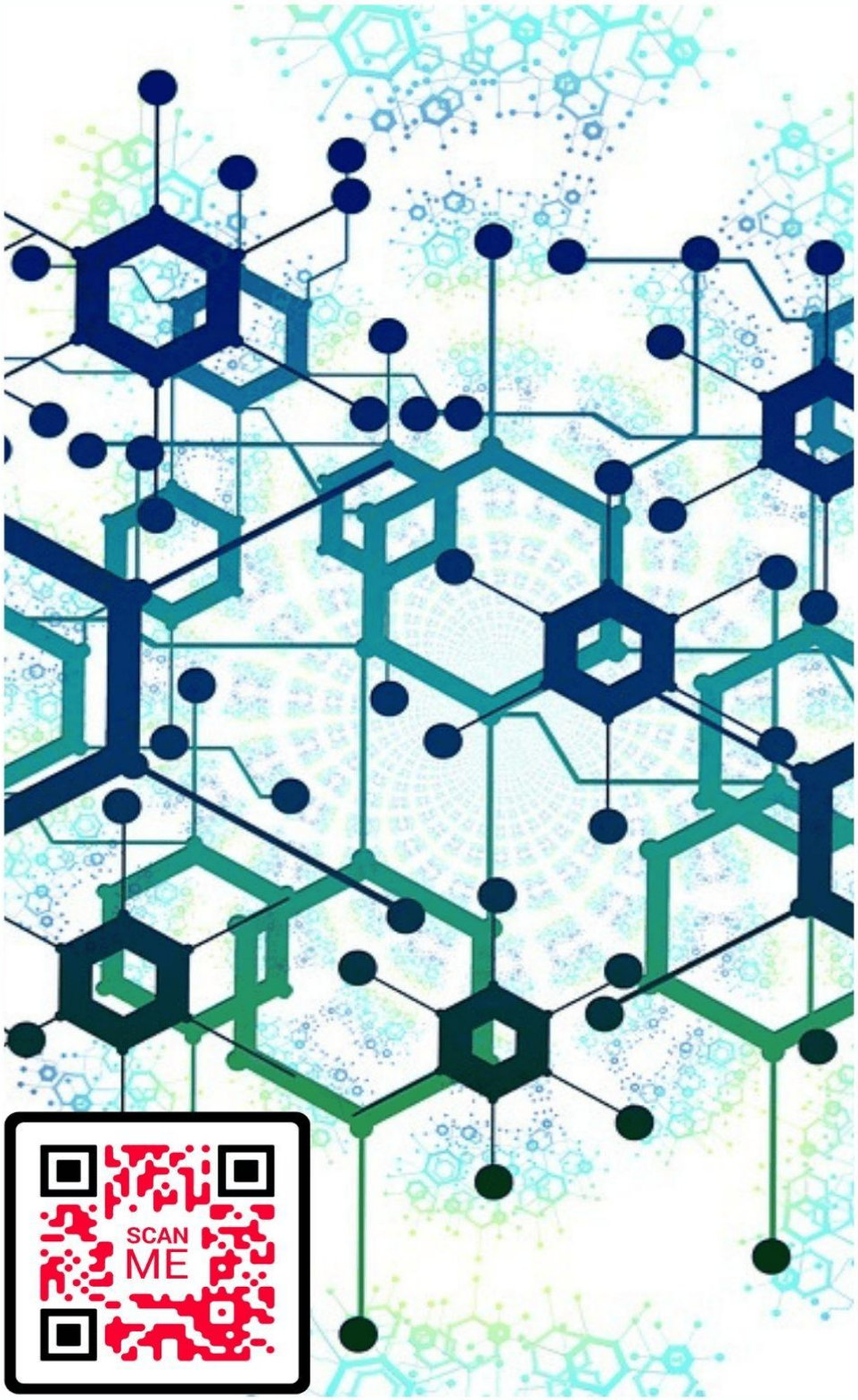
Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

ISSN (ONLINE) 2598-9936



INDONESIAN JOURNAL OF INNOVATION STUDIES

PUBLISHED BY

UNIVERSITAS MUHAMMADIYAH SIDOARJO

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

Table Of Contents

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

EDITORIAL TEAM

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

Article information

Check this article update (crossmark)

Check this article impact (*)

Save this article to Mendeley

(*) Time for indexing process is various, depends on indexing database platform

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

Indonesian Journal of Innovation Studies

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

Hyperparameter Optimization of Light Gradient Boosting Machine for Microcirculation Detection Wearable Data

Optimasi Hyperparameter Light Gradient Boosting Machine untuk Deteksi Mikrosirkulasi Data Wearable

Tatya Hanum Pramudita, tatya.hanum.2205356@students.um.ac.id, (1)
Program Studi Teknik Informatika, Universitas Negeri Malang, Indonesia

M. Zainal Arifin, arifin.mzainal@um.ac.id, ()
Program Studi Teknik Informatika, Universitas Negeri Malang, Indonesia

(1) Corresponding author

Abstract

General Background: Microcirculation disorders represent an early marker of chronic health conditions, yet existing detection approaches predominantly rely on invasive and resource-intensive procedures. **Specific Background:** Recent advances in wearable technology enable noninvasive microcirculation monitoring through Laser Doppler Flowmetry and Fluorescence Spectroscopy signals, which generate complex, nonstationary, and high-dimensional data that challenge conventional analytical methods. **Knowledge Gap:** Despite the proven capability of Light Gradient Boosting Machine models for wearable physiological data, limited studies have systematically combined feature selection, Bayesian hyperparameter optimization, and cohort-based validation for microcirculation condition detection using LDF-FS data. **Aims:** This study aims to optimize LightGBM performance for microcirculation condition detection by integrating feature importance–based selection and Bayesian hyperparameter tuning within a Stratified Group K-Fold validation framework. **Results:** Feature dimensionality was reduced from 34 to 22 informative variables, resulting in improved classification performance, with the optimized model achieving a ROC-AUC of 0.8632, accuracy of 88.04%, and recall of 80.00%. SHAP-based analysis identified age, body mass index, and skin temperature as dominant physiological predictors. **Novelty:** The study presents an integrated optimization pipeline combining feature selection, Bayesian optimization, and subject-level validation on wearable LDF-FS data. **Implications:** The findings support the potential of optimized LightGBM models as interpretable and reliable components of noninvasive wearable-based microcirculation monitoring systems.

Highlights

- Feature selection reduced dimensionality while maintaining robust classification performance
- Bayesian optimization improved sensitivity in detecting microcirculation conditions
- SHAP analysis revealed dominant demographic and physiological predictors

Keywords

Bayesian Optimization; Microcirculation Detection; Feature Selection; LightGBM; Wearable LDF-FS

Published date: 2026-01-11

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

I. Pendahuluan

Mikrosirkulasi darah merupakan bagian penting dari sistem peredaran darah yang bertanggung jawab terhadap distribusi oksigen, nutrisi, serta pembuangan sisa metabolisme pada jaringan. Sistem ini mencakup arteriola, kapiler, dan venula dengan ukuran yang sangat kecil, namun memiliki fungsi vital dalam menjaga homeostasis tubuh. Regulasi mikrosirkulasi berlangsung melalui interaksi kompleks antara faktor endotelial, neurogenik, miogenik, serta sinyal metabolik yang bekerja secara dinamis menyesuaikan kebutuhan jaringan. Ketika terjadi gangguan pada mikrosirkulasi, jaringan tidak dapat memperoleh suplai oksigen dan nutrisi yang memadai sehingga memicu kerusakan seluler. Kondisi ini menjadi salah satu faktor utama dalam perkembangan berbagai penyakit kronis, termasuk hipertensi, diabetes mellitus, penyakit kardiovaskular, hingga komplikasi pasca infeksi seperti COVID-19 yang terbukti mengubah pola aliran darah mikro [1].

Deteksi dini perubahan fungsi mikrosirkulasi sangat penting dilakukan sebagai langkah preventif untuk mencegah komplikasi serius. Metode konvensional umumnya bersifat invasive, berbiaya tinggi, memerlukan fasilitas laboraturium khusus, serta tidak mendukung pemantauan harian berkelanjutan. Keterbatasan ini mendorong pengembangan teknologi noninvasif berbasis perangkat wearable yang memungkinkan pemantauan fisiologis secara portable, real time dan kontinu. Salah satu pendekatan yang berkembang adalah integrasi Laser Doppler Flowmentry (LDF) dan Fluerescence Spectroscopy (FS). LDF digunakan untuk mengukur perfusi jaringan melalui pergeseran frekuensi cahaya akibat pergerakan sel darah merah. Sedangkan FS merekam fluoresensi biomolekul seperti NADH dan FAD yang mencerminkan status metabolisme sel. Kombinasi keduanya, yang dikenal sebagai LDF-FS memungkinkan analisis multimodal yang lebih komprehensif terhadap kondisi mikrosirkulasi [2].

Namun, data fisiologis yang dihasilkan bersifat komplek dan menantang untuk dianalisis. Pertama, sinyal bersifat non-stasioner, artinya nilai yang diperoleh dapat berubah dengan cepat mengikuti kondisi fisiologis tubuh, misalnya akibat aktivitas fisik, perubahan suhu, atau posisi tubuh. Kedua, sinyal sangat rentan terhadap noise dan artefak gerakan, misalnya akibat pergeseran sensor atau getaran tubuh, yang dapat menyebabkan distorsi. Ketiga, terdapat variabilitas tinggi antar individu yang dipengaruhi oleh faktor anatomi kulit, pigmen, ketebalan jaringan, serta gaya hidup.Keempat, sinyal LDF-FS memiliki struktur spectral yang rumit karena terdiri dari berbagai komponen osilasi fisiologis yang muncul pada rentang frekuensi berbeda, seperti osilasi endotelial, neurogenic, miogenik, respirasi, dan kardiak [2]. Kompleksitas data menyebabkan metode analisis konvensional kurang mampu menangkap pola laten secara akurat. Oleh karena itum pendekatan machine learning menjadi alternatif yang efektif karena mampu mengektraksi pola kompleks dari data berdimensi tinggi. Salah satu algoritma yang banyak digunakan adalah Light Gradiend Boosting Machine (LightGBM), pengembangan dari gradient boosting decision tree yang mengutamakan efisiensi dan akurasi melalui strategi pertumbuhan pohon leaf wise serta optimasi memori. Dibandingkan algoritma lain seperti Random Forest, Support Vector Machine, dan CatBoost, LightGBM lebih cepat, efisien, dan andal dalam menangani data dengan jumlah fitur besar [3]. Studi sebelumnya juga membuktikan efektivitas LightGBM pada data medis berbasis wearable, seperti penelitian Nguyen et al (2025) yang menggunakan dataset multimodal LDF-FS dari 132 partisipan di 19 negara untuk mendeteksi tingkat stress berdasarkan skor DAS-21 [4].

Dataset tersebut bersofat open access untuk keperluan riset non komersial, sehingga valid digunakan dalam penelitian lanjutan. Hasil penelitian menunjukkan bahwa LightGBM dengan seleksi 10 fitur terpenting dan validasi menggunakan 5 fold cross validation menghasilkan nilai ROD AUC sebesar 0.7168 pada tugas klasifikasi biner, lebih unggul dibandingkan algoritma lain seperti Random Forest, SVM. Dan CatBoost. Hasil ini menunjukkan potensi LightGBM sebagai algoritma andalan dalam klasifikasi kondisi kesehatan berbasis sinyal fisiologis kompleks.

Kinerja LightGBM sangat ditentukan oleh pengaturan hiperparameter seperti learning rate, number of leaves, max depth, feature fraction, dan bagging fraction yang memepngaruhi kompleksitas, generalisasi, dan stabilitas midel. Konfigurasi yang terlalu sederhana dapat menyebabkan underfitting, sedangkan pengaturan yang terlalu kompleks berisiko menimbulkan overfitting, sehingga kemampuan generalisasi menurun, terutama pada data fisiologis yang kompleks seperti sinyal LDF-FS. Oleh karena itu, hyperparameter tuning menjadi langkah krusial untuk memperoleh konfigurasi optimal yang menyeimbangkan bias dan varians model. Dalam lima tahun terakhir, berbagai metode optimasi hyperparameter telah dikembangkan.metode konvensional seperti Grid Search dan Ramdom Search masih banyak digunakan, namun cenderung memakan waktu karena melakukan eksplorasi yang luas terhadap ruang parameter. Metode modern seperti Bayesian Optimization tetap menjadi terdepan dalam optimasi hyperparameter karena menggunakan pendekatan probabilistic untuk memperkirakan kombinasi parameter terbaik berdasarkan percobaan sebelumnya, dengan eifisiensi sampel yang jauh lebih tinggi daripada pencarian acak dan grid search.

Selain itu, terdapat pula pendekatan berbasis AutoML (Automated Machine Learning) yang semakin populer karena dapat menggabungkan hyperparameter tuning, pemilihan model, serta preprocessing secara otomatis dengan hasil yang kompetitif [5]. Studi terbaru juga mulai mengintegrasikan metode optimasi berbasis metaheuristic, seperti Particle Swarm Optimization (PSO) dan Genetic Algorithms (GA), untuk mencari konfigurasi parameter LightGBM yang optimal dalam permasalahan medis [6].

Selain hiperparameter, seleksi fitur berperan penting dalam meningkatkan kinerja model dengan memilih fitur paling relevan dan menghilangkan fitur yang kurang berkontribuso. Tahapa ini sangat krusial pada data LDF-FS yang berdimensi tinggi dan mencakup fitur domain waktu, frekuensi serta nonlinier. Tanpa seleksi fitur, model cenderung menjadi terlalu kompleks, rentan terhadap overvitting dan sulit diinterpretasikan. Dalam konteks data fisiologis, feature selection juga membantu mengidentifikasi biomarker potensial yang memiliki makna bilogis, misalnya osilasi endotelial atau rasio fluoresensi tertentu yang berkaitan dengan metabolisme seluler.

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

Penelitian terkini menunjukkan bahwa integrasi feature selection dengan LightGBM dapat meningkatkan kinerja sekaligus memberikan model yang lebih interpretative. Guldogan dan Yagin (2025) membuktikan bahwa penggunaan metode seleksi fitur multi objektif berbasis SHAP dalam LightGBM pada analisis metabolomik kanker payudara menghasilkan peningkatan akurasi serta kemampuan menjelaskan kontribusi fitur terhadap prediksi [7].

Noorozy et al. (2023) melaporkan bahwa kombinasi seleksi fitur dan algoritma boosting seperti LightGBM secara signifikan meningkatkan stabilitas model pada prediksi penyakit jantung [8]. Oleh karena itu, seleksi fitur tidak hanya berfungsi untuk reduksi dimensi, tetapi juga meningkatkan transparansi model dalam aplikasi medis. Berdasarkan hal tersebut, penelitian ini berfokus pada optimasi LightGBM melalui hyperparameter tuning dan seleksi untuk mendeteksi kondisi mikrosirkulasi berbasis data wearable LDF-FS. Berbeda dengan penelitian sebelumnya, penelitian ini tidak hanya menggunakan LightGBM untuk klasifikasi, tetapi juga menerapkan kombinasi tahap hyperparameter tuning dan feature selection secara berurutan untuk memperoleh konfigurasi dan subset fitur yang optimal. Dengan demikian, penelitian ini tidak meneliti interaksi matematis antar keduanya, tetapi menilai pengaruh gabungan tahap optimasi tersebut terhadap performa model secara keseluruhan.

Maka dari itu, penelitian ini bertujuan menghasilkan model yang tidak hanya akurat dan efisien, tetapi juga interpretative, khususnya untuk aplikasi medis. Interpretabilitas penting untuk memahami dasar pengambilan keputusan model, sehingga pendekatan Explainable *AI* (XAI) seperti SHAP digunakan guna mengindentifikasi kontribusi setiap fitur terhadap prediksi. Pada data fisiologis, interpretasi ini membantu menungkap biomarker yang relevan secara klinis, seperti osilasi endotelial dan parameter fluoresensi yang berkaitan dengan kondisi mikrosirkulasi. Dengan demikian, diharapkan hasil penelitian ini dapat mendukung pengembangan teknologi wearable sebagai perangkat pemantauan kesehatan yang bersifat preventif, non-invasif, dan dapat digunakan secara real-time.

II. Metode

A. Kerangka Penelitian

Penelitian ini dirancang menggunakan kerangka penelitian yang tersusun secara sistematis untuk menggambarkan alur penelitian dari tahap awal hingga akhir. Tahapan penelitian meliputi pengumpulan dataset, pembagian data, preprocessing data, seleksi fitur, pembangunan model klasifikasi, evaluasi model, serta analisis interpretabilitas model.

Dataset yang digunakan merupakan dataset publik LDF–FS yang diperkenalkan oleh Nguyen et al. (2025), yang memuat data sinyal mikrosirkulasi hasil pengukuran wearable berbasis Laser Doppler Flowmetry (LDF) dan Fluorescence Spectroscopy (FS), serta hasil kuesioner DASS-21 sebagai label kondisi psikologis [4]. Data kemudian diproses melalui tahapan preprocessing, seleksi fitur menggunakan feature importance LightGBM, serta klasifikasi menggunakan algoritma Light Gradient Boosting Machine (LightGBM). Evaluasi dan interpretasi model dilakukan menggunakan metrik klasifikasi dan metode SHAP. Kerangka penelitian ditunjukkan pada Gambar 1.

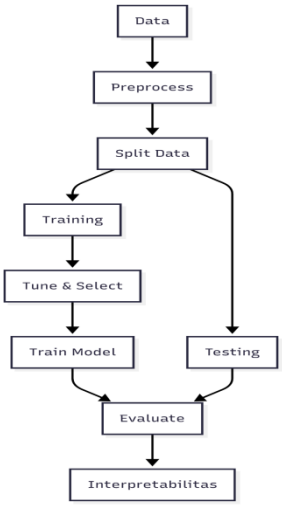


Figure 1. *Kerangka Penelitian*

B. Sumber Data

Penelitian ini menggunakan data sekunder berupa dataset wearable LDF–FS yang diperoleh dari penelitian terdahulu dengan lisensi penelitian terbuka. Dataset melibatkan 132 partisipan dari 19 negara dengan rentang usia 18–94 tahun. Data yang digunakan mencakup sinyal perfusi mikrosirkulasi, fluoresensi NADH/FAD, parameter fisiologis (suhu tubuh, detak jantung, BMI), metadata partisipan, serta hasil kuesioner DASS-21.

Dataset terdiri dari 41 fitur, yang mencakup fitur numerik dan kategorikal, dengan label akhir berupa kondisi Wellbeing dan Non-Wellbeing berdasarkan indikator stres, kecemasan, dan depresi. Dataset disajikan dalam format CSV dan Excel, serta

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

memungkinkan replikasi penelitian secara akurat.

C. Preprocessing Data

Preprocessing data dilakukan untuk memastikan kualitas dan kesiapan dataset sebelum pemodelan. Tahapan preprocessing meliputi:

1. Data Cleaning

Tahap data cleaning bertujuan meningkatkan kualitas dataset dengan menangani nilai hilang, inkonsistensi, dan nilai tidak wajar yang berpotensi menimbulkan bias serta menurunkan performa model klasifikasi. Proses ini dilakukan dengan mengidentifikasi atribut yang memiliki missing values dan nilai ekstrem pada data fisiologis. Baris data yang tidak informatif dihapus, sedangkan atribut numerik dengan jumlah nilai hilang yang kecil diimputasi menggunakan metode K-Nearest Neighbors (KNN). Metode ini dipilih karena mampu memperkirakan nilai hilang berdasarkan kemiripan antar sampel, sehingga mempertahankan distribusi data serta hubungan antar fitur fisiologis LDF–FS secara lebih akurat dibandingkan metode imputasi sederhana. Selain itu, dilakukan pemeriksaan terhadap nilai ekstrem atau tidak logis, seperti detak jantung di luar rentang fisiologis wajar, untuk menentukan apakah data tersebut perlu diperbaiki atau dihapus. Seluruh proses pembersihan data dilakukan secara sistematis guna memastikan dataset yang digunakan pada tahap pemodelan bersih, konsisten, dan representatif terhadap kondisi fisiologis partisipan [9].

2. Normalization

Normalisasi dilakukan untuk menyeragamkan skala atribut numerik agar perbedaan rentang nilai tidak menimbulkan bias pada proses pelatihan model. Pada penelitian ini, atribut numerik seperti Heart Rate dan fitur hasil ekstraksi sinyal fisiologis LDF–FS dinormalisasi menggunakan metode Min–Max Normalization ke rentang [0,1]. Normalisasi diterapkan setelah tahap data cleaning dan imputasi, sehingga data yang digunakan telah bebas dari nilai hilang dan kesalahan pencatatan. Hasil normalisasi menghasilkan dataset dengan skala yang seimbang dan siap digunakan pada tahap seleksi fitur dan klasifikasi menggunakan algoritma LightGBM.

3. Encoding Data Kategorikal

Tahap encoding dilakukan untuk memastikan atribut kategorikal dapat diproses dalam pemodelan machine learning tanpa menghilangkan makna kategorinya. Dalam penelitian ini, atribut kategorikal seperti tingkat tekanan darag dan status psikologis DASS-21 diproses menggunakan kemampuan bawaan LightGBM dalam menangani variable kategorikal dengan mengonversi tipe data ke format category. Oleh karena itu, proses encoding manual tidak dilakukan, kecuali untuk keperluan analisis tambahan atau penerapan algoritma pembanding. Pendekatan ini memungkinkan pemodelan yang lebih eifisien sekaligus mempertahakan informasi kategorikal secara optimal.

4. Data Splitting

Pembagian data dilakukan untuk mengevaluasi kemampuan generalisasi model terhadap data baru serta mencegah bias evaluasi. Pada penelitian ini digunakan metode Stratified Group K-Fold Cross-Validation (k = 5), yang menjaga keseimbangan distribusi kelas sekaligus memastikan seluruh sampel dari individu yang sama hanya berada pada satu fold. Pendekatan ini efektif mencegah data leakage pada dataset fisiologis LDF–FS yang memiliki lebih dari satu sampel per partisipan, sehingga performa model yang diperoleh merepresentasikan kemampuan prediksi pada individu yang belum pernah dilihat sebelumnya.

D. Proses Seleksi Fitur dengan Feature Importance

Seleksi fitur dilakukan untuk menentukan fitur yang paling relevan terhadap target klasifikasi sehingga model menjadi lebih efisien, stabil, dan mudah diinterpretasikan. Pada penelitian ini, seleksi fitur dilakukan setelah tahap preprocessing menggunakan pendekatan feature importance dari algoritma Light Gradient Boosting Machine (LightGBM) berdasarkan metrik gain, yang merepresentasikan kontribusi fitur dalam menurunkan fungsi loss. Proses seleksi fitur dilakukan di dalam skema Stratified Group K-Fold Cross-Validation untuk menjaga keseimbangan kelas sekaligus mencegah kebocoran data antar individu. Fitur diurutkan berdasarkan nilai importance dan dipilih menggunakan pendekatan Top-N features, kemudian dievaluasi berdasarkan stabilitas performa model menggunakan metrik ROC-AUC, precision, recall, dan F1-score.

Hasil seleksi menunjukkan bahwa dari seluruh fitur awal, diperoleh 22 fitur terbaik yang mampu mempertahankan performa klasifikasi sekaligus mengurangi kompleksitas dan beban komputasi model. Model selanjutnya dilatih menggunakan fitur terpilih dan dianalisis menggunakan metode SHAP untuk mendukung interpretabilitas hasil prediksi. Validasi ahli dilakukan untuk memastikan bahwa fitur terpilih relevan secara praktis dan klinis.

E. Proses Klasifikasi dengan Algoritma LightGBM

Proses klasifikasi dilakukan menggunakan algoritma Light Gradient Boosting Machine (LightGBM) pada dataset yang telah melalui tahap preprocessing dan seleksi fitur. LightGBM dipilih karena eifisnesi komputasi, kemampuannya menangani data berdimensi tinggi, serta efektivitasnya terhadap ketidakseimbangan kelas. Validasi dilakukan menggunakan Stratified Group K-Fold (k=5) untuk menjaga distribusi kelas dan mencegah kebocoran data antar subjek. Proses klasifikasi dimulai dari

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

model baseline dengan parameter default, kemudian dioptimalkan melalui Bayesian Optimization dengan penerapan eraly stopping.

Konfigurasi hyperparameter optimal digunakan untuk melatih model akhir, yang kemudian dievaluasi menggunakan metrik accuracy, precision, recall, F1-score, ROC-AUC, dan PR-AUC. Model akhir beserta hasil feature importance selanjutnya dianalisis menggunakan metode SHAP untuk mendukung interpretabilitas hasil prediksi. Pendekatan ini memastikan bahwa model yang dihasilkan tidak hanya akurat, tetapi juga dapat dijelaskan secara ilmiah.

1. Pembangunan Model Baseline

Tahap awal klasifikasi dilakukan dengan membangun model baseline menggunakan algoritma Light Gradient Boosting Machine (LightGBM) dan seluruh fitur hasil seleksi, tanpa optimasi hyperparameter. Model baseline ini digunakan sebagai acuan awal untuk menilai peningkatan performa setelah proses optimasi dilakukan. Pembagian data dilakukan menggunakan Stratified Group K-Fold Cross-Validation (k = 5) guna menjaga keseimbangan distribusi kelas Wellbeing dan Non-Wellbeing, sekaligus memastikan bahwa data dari partisipan yang sama tidak muncul pada data latih dan data validasi secara bersamaan sehingga risiko kebocoran data dapat dihindari. Performa awal model dievaluasi menggunakan metrik accuracy, precision, recall, F1-score, ROC-AUC, dan PR-AUC, dengan nilai rata-rata dari seluruh fold sebagai representasi performa baseline.

2. Optimasi Hyperparameter

Optimasi hyperparameter dilakukan untuk memperoleh konfigurasi parameter terbaik yang mampu meningkatkan performa model dan mengurangi risiko overfitting. Proses optimasi menggunakan pendekatan Bayesian Optimization yang dikombinasikan dengan Stratified Group K-Fold Cross-Validation (k = 5) pada data latih.

Pendekatan ini dipilih karena mampu mencari konfigurasi parameter optimal secara adaptif dan efisien. Optimasi mencakup kompleksitas model, learning rate, pembobotan kelas, dan regularisasi, dengan evaluasi kinerja menggunakan ROD-AUC sebagai fungsi objektif serta penerapan early stopping saat peningkatan validasi tidak signifikan.

3. Pelatihan dan Evaluasi Model Akhir

Model akhir dilatih menggunakan konfigurasi hyperparameter optimal yang diperoleh dari proses optimasi sebelumnya dan dievaluasi secara objektif menggunakan skema Stratified Group K-Fold Cross-Validation (k = 5). Model mengklasifikasikan data ke dalam dua kelas, yaitu Wellbeing dan Non-Wellbeing, di mana deteksi kelas Non-Wellbeing menjadi prioritas utama. Evaluasi performa dilakukan menggunakan accuracy, precision, recall, F1-score, ROC-AUC, dan PR-AUC, serta dianalisis menggunakan confusion matrix. Selain evaluasi kuantitatif, dilakukan analisis interpretabilitas menggunakan metode SHAP untuk menjelaskan kontribusi fitur terhadap hasil prediksi, sehingga model yang dihasilkan tidak hanya memiliki performa yang baik tetapi juga dapat dipertanggungjawabkan secara ilmiah.

F. Analisis Feature Importance dan Interpretabilitas

Analisis feature importance dan interpretabilitas dilakukan untuk memahami kontribusi masing-masing fitur terhadap hasil klasifikasi serta memastikan bahwa model yang dibangun tidak hanya memiliki performa yang baik, tetapi juga dapat dijelaskan secara fisiologis dan psikologis. Tahap ini bertujuan memastikan keputusan model memiliki dasar ilmiah yang kuat pada anlisis sinyal fisiologis. Feature importance dihitung dari LightGBM menggunakan metric gain untuk merepresentasikan kontribusi fitur terhadap penurunan loss. Hasil seleksi menghasilkan 22 fitur paling signifikan yang diurutkan berdasarkan tingkat kepentingannya dan digunakan dalam model akhir. Selanjutnyam interpretabilitas dianalisis menggunakan SHAP untuk menjelaskan arah dan besaran kontribusi fitur terhadap prediksi kelas Wellbeing dan Non Wellbeing, yang divisualisasikan melalui summary plot dan dependence plot. Pendekatan ini mendukung evaluasi teknis sekaligus interpretasi ilmiah terhadap hubungan antara fitur fisiologis dan kondisi psikologis partisipan, yang selanjutnya dibahas pada bagian pembahasan [10].

III. Hasil dan Pembahasan

A. Hasil Preprocessing Data

Adanya proses preprocessing data dalam penelitian ini bertujuan untuk menyiapkan dataset agar sesuai dengan kebutuhan pemodelan klasifikasi menggunakan algoritma Light Gradient Boosting Machine (LightGBM). Dataset penelitian ini berasal dari sensor wearable Laser Doppler Flowmentry (LDF) dan Fluorescence Spectroscopy (FS), serta kuesioner psikologis DASS-21. Data LDF-FS bersifat kompleks, nonstasioner, dan berpotensi mengandung noise, sedangkan data DASS-21 bersifat kategorikal dan rentan terhadap ketidakkonsistenan. Oleh karena itu, dilakukan beberapa tahapan preprocessing yang meliputi data cleaning, normalisasi, dan encoding variabel kategorikal. Hasil dari setiap tahapan preprocessing tersebut dijelaskan sebagai berikut.

1. Data Cleaning

Tahapan data cleaning dilakukan untuk memastikan kualitas dataset sebelum digunakan dalam proses pemodelan. Dataset awal terdiri dari 460 record data dengan 42 atribut, yang kemudian direduksi menjadi 34 atribut setelah penghapusan kolom metadata yang tidak relevan, seperti identitas pasien, nomor urut, serta variabel yang berpotensi menyebabkan data leakage

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

terhadap variabel target.

Berdasarkan hasil inspeksi data, tidak ditemukan baris data bertipe filter noise, sehingga tidak ada record yang dihapus pada tahap ini. Namun demikian, ditemukan missing value pada 5 atribut numerik, dengan total 25 baris data yang mengandung nilai kosong. Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan metode K-Nearest Neighbors (KNN) Imputation dengan jumlah tetangga (k) sebanyak 5. Metode ini memperkirakan nilai yang hilang berdasarkan kemiripan pola antar fitur numerik lainnya, sehingga mampu mempertahankan struktur dan distribusi alami data fisiologis [11].

Pemilihan metode KNN Imputation didasarkan pada kemampuannya dalam menangkap hubungan antar fitur yang kompleks, khususnya pada data sensor LDF–FS, dibandingkan metode imputasi sederhana seperti mean atau median yang berpotensi mengaburkan variasi fisiologis. Setelah proses imputasi dilakukan, seluruh dataset dinyatakan lengkap tanpa missing value dan siap untuk diproses pada tahap selanjutnya.

	Jumlah missing values	**Jumlah atribut**
Sebelum Data Cleaning	25	42
Setelah Data Cleaning	0	34

Table 1. *Kondisi Dataset Sebelum dan Sesudah Data Cleaning*

2. Normalisasi

Tahap berikutnya adalah normalisasi data, yang bertujuan untuk menyeragamkan skala nilai antar atribut numerik agar tidak terjadi dominasi fitur tertentu dalam proses pembelajaran model. Normalisasi dilakukan hanya pada atribut numerik menggunakan metode Min–Max Normalization, dengan mentransformasikan nilai fitur ke dalam rentang 0 hingga 1.

Proses normalisasi dilakukan dengan skema fit pada data latih (training set) dan kemudian diterapkan (transform) pada data uji (test set). Pendekatan ini bertujuan untuk mencegah terjadinya data leakage serta memastikan bahwa proses evaluasi model merefleksikan kondisi data yang sesungguhnya. Tabel 2 menampilkan contoh hasil normalisasi data menggunakan metode Min–Max Normalization.

No	**Kv**	**M**	**σ**	**T**	**A365**	**A460**	**Anadn**
1	0,21	0,34	0,18	0,52	0,48	0,41	0,32
2	0,36	0,35	0,42	0,53	0,56	0,19	0,21
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
459	0,44	0,39	0,58	0,61	0,22	0,49	0,39
460	0,12	0,46	0,16	0,57	0,19	0,36	0,57

Table 2. *Hasil Proses Normalisasi*

Hasil normalisasi menunjukkan bahwa seluruh atribut numerik berhasil diseragamkan dalam rentang $0 \leq x \leq 1$ tanpa mengubah hubungan proporsional antar variabel. Nilai pada tabel merupakan contoh hasil transformasi menggunakan metode Min–Max Normalization yang memetakan setiap atribut numerik ke dalam rentang 0 hingga 1. Normalisasi dilakukan menggunakan parameter yang dipelajari dari data latih dan diterapkan pada data uji. Nilai yang ditampilkan bersifat representatif dan tidak mencerminkan keseluruhan dataset.

Metode Min–Max Normalization dipilih karena algoritma LightGBM tidak memerlukan asumsi distribusi data tertentu, serta data fisiologis pada penelitian ini memiliki rentang nilai yang relatif terkontrol sehingga tidak memerlukan standardisasi berbasis distribusi seperti Z-Score [12]. Setelah tahap normalisasi dilakukan, data dinyatakan siap untuk diproses pada tahap encoding dan seleksi fitur.

3. Encoding Variabel Kategorikal

Sebagian besar algoritma pembelajaran mesin, termasuk LightGBM, memerlukan input data dalam format numerik. Dalam dataset penelitian ini, terdapat sejumlah variabel independen yang bersifat kategorikal (object), antara lain Gender, Ethnicity,

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

Disease, Smoking, dan Blood Pressure. Variabel-variabel ini memuat informasi kualitatif yang perlu ditransformasikan agar dapat diproses secara matematis.

Proses tranformasi dilakukan menggunakan metode Label Encoding. Metode ini mengonversi setiap label kategori unik dalam suatu fitur menjadi bilangat bulat (integer) secara berurutan berdasarkan urutan abjad (alfabites). Penerapan Label Encoding dilakukan secara interatif pada seluruh kolam kategorikal yang terdeteksi dalam dataframe. Metode ini dipilih karena efisiensi memori yang lebih baik dibandingkan One Hot Encoding. Sebagai contoh, variable Ethnicity dan Disease memiliki kardinalitas (jumlah variasi unik) yang tinggi. Jika menggunakan One-Hot Encoding, dimensi data akan membengkak signifikan. Dengan Label Encoding, struktur data tetap ringkas tanpa mengorbankan informasi kategori [13]. Tabel 3 berikut menampilkan sampel hasil pemetaan kategori ke nilai numerik berdasarkan output program.

Nama Variabel	Kategori	Hasil Encoding
Gender	Female	0
	Male	1
Ethnicity	India	4
	Pakistan	15
	Vietnam	19
Disease	Anemia	0
	Diabetes Hypertension	2
	Hypertension	14
	No	20
Blood Pressure	Elevated	0
	High Blood	1
	Hypertensive Crisis	2
	Normal	3

Table 3. *Hasil Encoding Variabel Kategorikal*

Nilai numerik di atas merupakan hasil pemetaan otomatis oleh LabelEncoder. Angka yang dihasilkan berfungsi sebagai pembeda identitas kategori (nominal), bukan menunjukkan tingkatan atau urutan derajat (ordinal). Setelah tahap ini, seluruh fitur dalam dataset telah memiliki format numerik yang seragam dan siap untuk diproses ke tahap normalisasi serta pembagian data latih dan uji.

B. Data Splitting

Setelah seluruh proses preprocessing selesai, tahap selanjutnya adalah pembagian data (data splitting) untuk keperluan pelatihan dan evaluasi model. Dalam penelitian ini diterapkan strategi Stratified Group K-Fold Cross-Validation dengan jumlah lipatan (fold) sebanyak lima. Pemilihan metode ini didasarkan pada karakteristik dataset wearable, di mana satu subjek (pasien) memiliki beberapa baris data rekaman yang berbeda (data time-series atau hasil segmentasi). Dataset ini

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

terdiri dari total 460 baris data (sampel) yang berasal dari 200 identitas subjek unik. Jika menggunakan metode random split, terdapat risiko terjadinya data leakage, yaitu kondisi di mana sampel dari subjek yang sama muncul di data latih dan data uji secara bersamaan. Hal ini dapat menyebabkan model tampak memiliki akurasi tinggi, bukan karena mampu mengenali pola penyakit secara umum, melainkan karena mengenali karakteristik subjek tertentu.

Dengan penerapan Stratified Group K-Fold Cross-Validation, pembagian data dilakukan dengan memperhatikan dua aspek penting. Pertama, grouping yaitu seluruh sampel dari satu subjek ditempatkan hanya pada satu subset, sehingga data dari subjek yang sama tidak tersebar di training set maupun validation set. Kedua, stratification, yang menjaga proporsi distribusi kelas target tetap seimbang antara training set dan validation set pada setiap iterasi [14]. Dengan mekanisme ini, setiap subset data tetap representatif terhadap distribusi kelas target, sekaligus mencegah kebocoran informasi dari subjek yang sama.

Skema pembagian membagi dataset menjadi lima bagian, di mana pada setiap iterasi empat bagian digunakan untuk melatih model dan satu bagian digunakan untuk validasi. Proses ini diulang sebanyak lima kali sehingga setiap bagian data pernah menjadi data validasi. Tabel 4 menunjukkan distribusi rata-rata jumlah subjek dan sampel pada skema 5-Fold yang diterapkan. Mekanisme pembagian ini memastikan bahwa evaluasi model bersifat objektif dan mencerminkan kemampuan generalisasi terhadap subjek baru (unseen subjects).

Subset Data	Rata-Rata Jumlah Subjek	Rata-Rata Jumlah Sampel
Training Set	160	368
Validation Set	40	92
Total	200	450

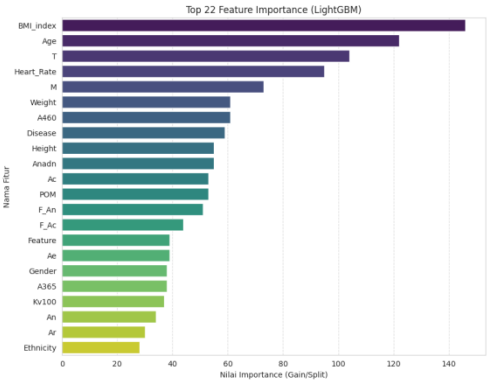
Table 4. *Distribusi Data pada Skema Stratified Group K-Fold*

Angka jumlah subjek dan sampel pada tabel merupakan rata-rata per fold, karena terdapat sedikit variasi jumlah record per individu. Metode Stratified Group K-Fold menjamin tidak ada irisan subjek antara training set dan validation set, sehingga evaluasi model bebas dari data leakage [15]. Dengan mekanisme pembagian ini, hasil evaluasi model diharapkan bersifat objektif (unbiased) dan mencerminkan kemampuan generalisasi model terhadap subjek baru yang belum pernah dilihat sebelumnya (unseen subjects).

C. Hasil Seleksi Fitur dengan LightGBM

Setelah dataset terbagi menjadi data latih dan data uji, tahap selanjutnya adalah seleksi fitur (feature selection). Tahap ini bertujuan memilih fitur paling berpengaruh sekaligus mereduksi dimensi data dengan menghilangkan fitur yang tidak relevan. Tanpa seleksi yang tepat, penggunaan fitur berlebih dapat meningkatkan kompleksitas komputasi dan risiko overfitting. Oleh karena itu, penelitian ini menggunakan seleksi fitur bawaan (embedded method) dan algoritma LightGBM sebagai alternatif metode konvensional. Algoritma ini secara otomatis menghitung skor kepentingan fitur (feature importance score) selama proses pelatihan pada baseline model. Metrik yang digunakan adalah Total Gain, yang mengukur seberapa besar penurunan loss (ketidakmurnian) yang disumbangkan oleh suatu fitur setiap kali fitur tersebut digunakan sebagai titik pemecahan (split point) dalam pohon keputusan [16].

Berdasarkan hasil eksperimen, ditetapkan ambang batas untuk mengambil 22 fitur dengan skor tertinggi (Top-22 Features) yang akan digunakan pada tahap pemodelan akhir. Visualisasi tingkat kepentingan fitur disajikan pada Gambar 2 berikut.



Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

Figure 2. *Visualisasi Top-22 Feature Importance LightGBM*

Berdasarkan Gambar 2, terlihat jelas urutan fitur yang paling berpengaruh terhadap keputusan model. Fitur yang berada di posisi teratas grafik batang memiliki nilai importance yang jauh lebih tinggi dibandingkan fitur-fitur di bawahnya. Hal ini mengindikasikan bahwa fitur-fitur tersebut, terutama yang berkaitan dengan karakteristik fisik dan parameter sensor, merupakan indikator utama dalam membedakan kelas target.

Sebanyak 22 fitur terpilih ini kemudian digunakan sebagai input eksklusif untuk pelatihan model utama. Reduksi fitur menjadi 22 atribut ini terbukti mampu menjaga efisiensi komputasi sekaligus membuang noise yang dibawa oleh fitur-fitur dengan skor kepentingan rendah. Rincian skor untuk fitur-fitur terpilih ditampilkan pada Tabel 4.5.

No	Nama Fitur	Skor
1	BMI_index	146.0000
2	Age	122.0000
3	T	104.0000
4	Heart_Rate	95.0000
5	M	73.0000
6	Weight	61.0000
7	A460	61.0000
8	Disease	59.0000
9	Height	55.0000
10	Anadn	55.0000
11	Ac	53.0000
12	POM	53.0000
13	F_An	51.0000
14	F_Ac	44.0000
15	Feature	39.0000
16	Ae	39.0000
17	Gender	38.0000

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

18	A365	38.0000
19	Kv100	37.0000
20	An	34.0000
21	Ar	30.0000
22	Ethnicity	28.0000

Table 5. *Daftar 22 Fitur Terpilih Berdasarkan Feature Importance*

Berdasarkan Tabel 5, terlihat pola yang menarik di mana kombinasi antara data antropometri dan data sinyal sensor mendominasi peringkat teratas. Fitur BMI INDEX (Skor: 146.0) dan Age (Skor:122.0) menempati dua posisi tertinggi, mengindikasikan bahwa massa tubuh dan faktor usia merupakan predictor fundamental dalam menentukan kondisi kesehatan mikrosirkulasi seseorang.

Selain faktor demografis, parameter fisiologis dari sensor LDF-FS seperti T (Skor: 104.0), Heart Rate (Skor; 95.0), dan parameter spectral M (Skor:73.0) juga memberikan kontribusi signifikan. Hal ini membuktikan bahwa model berhasil menangkan hubungan kompleks antara kondisi fisik dasar pasien dengan sinyal mikrosirkulasi yang direkam oleh alat. Sebaliknya, fitur seperti Ethnicity memiliki pengaruh yang relatif lebih rendah (Skor: 28.0), menunjukkan bahwa kondisi fisiologis lebih bersifat universal dibandingkan latar belakang etnis [17].

D. Hasil Model Baseline LightGBM

Pemodelan awal dilakukan dengan membangun model baseline menggunakan konfigurasi parameter bawaan (default hyperparameters) dari algoritma LightGBM. Tahap ini bertujuan untuk menetapkan titik acuan (benchmark) kinerja sebelum dilakukan optimasi lebih lanjut. Evaluasi model baseline dilaksanakan melalui dia scenario pengujian guna mengukur efektivitas reduksi dimensi. Scenario pertama (All Features) melibatkan pelatihan model menggunakan seluruh atribut yang tersedia, sedangkan sekanrio kedua (Selected Features) membatasi input model hanya pada 22 fitur terbaik hasil seleksi fitur. Skema validasi yang digunakan pada kedua scenario tetap konsisten, yaitu Stratified Group K-Fold Cross Validation (k=5).

1. Evaluasi Baseline pada Seluruh Fitur

Pada skenario pertama, model dilatih dengan memanfaatkan seluruh dimensi data tanpa pengurangan fitur. Hasil evaluasi rata-rata dari 5 fold validasi disajikan pada Tabel 4.6.

Metrik Evaluasi	Rata-rata Skor (Mean)	Standar Deviasi
ROC-AUC	0.6939	0.1468
Akurasi	0.7094	0.1068
Presisi	0.3427	0.2806
Recall	0.2963	0.2238
F1-Score	0.3160	0.2457

Table 6. *Hasil Evaluasi Model Baseline LightGBM dengan Semua Fitur*

Berdasarkan Tabel 6, model baseline dengan semua fitur menghasilkan performa yang moderat dengan skor ROC-AUC sebesar 0,6939. Nilai Recall yang rendah (0,2963) mengindikasikan bahwa model masih kesulitan mendeteksi kelas positif (kondisi sakit/stres) di tengah banyaknya fitur yang mungkin mengandung noise atau informasi yang tidak relevan.

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

2. Evaluasi Baseline pada 22 Fitur Terpilih

Skenario kedua menerapkan model pada dataset yang telah direduksi menjadi 22 fitur paling signifikan. Hasil evaluasi kinerjanya ditampilkan pada Tabel 7.

Metrik Evaluasi	Rata-rata Skor (Mean)	Standar Deviasi
ROC-AUC	0.7061	0.1701
Akurasi	0.7418	0.1266
Presisi	0.3823	0.3341
Recall	0.3228	0.3011
F1-Score	0.3488	0.3178

Table 7. *Hasil Evaluasi Model Baseline LightGBM dengan 22 Fitur Terpilih*

Berdasarkan perbandingan antara Tabel 6 dan Tabel 7, penerapan seleksi fitur terbukti memberikan dampak positif terhadap kinerja model.Skor ROC-AUC meningkat menjadi 0,7061 dan akurasi 74,18%, menandakan reduksi dimensi efektif menghilangkan fitur tidak relevan dan memperjelas pola data. Namun, rendahnya recall dan F2 score menunjukkan bias terhadap kelas mayoritas, sehingga diperlukan optimasi hiperpaeameter lanjutan untuk meningkatkan sensitivitas model.

E. Hasil Optimasi Hyperparameter dan Evaluasi Model Akhir

Evaluasi model baseline menunjukkan ketidakseimbangan kelas sebagai penyebab rendahnya recall. Untuk mengatasinya, diterapkan optimasi dua tahap, yakni Bayesian Optimization dengan focus pada bobot kelas serts threshold tuning untuk menyeimbangkan precision dan recall.

1. Implementasi Bayesian Optimization

Proses optimasi dilakukan untuk mencari kombinasi hyperparameter terbaik pada model LightGBM. Ruang pencarian (search space) dibatasi pada parameter yang mengontrol kompleksitas model (num_leaves, max_depth) dan regularisasi (lambda_l1, lambda_l2). Selain itu, parameter (scale_pos_weight) diatur secara dinamis dengan mengalikan rasio kelas dasar (Base Ratio) dengan faktor pengali (weight_mult) untuk memberikan penalti lebih besar pada kesalahan klasifikasi kelas minoritas.

Berdasarkan perhitungan pada data latih, diperoleh Base Class Ratio sebesar 3.49 yaitu perbandingan jumlah data sehat terhadap data sakit. Melalui 20 iterasi pencarian, algoritma Bayesian Optimization berhasil menemukan konfigurasi optimal dengan nilai pembobotan akhir (Final Weight) sebesar 4.81. Nilai ini mengindikasikan bahwa model memberikan perhatian hampir 5 kali lipat lebih besar terhadap kelas positif (sakit/stres) dibandingkan kelas negatif. Rincian ruang pencarian dan parameter terbaik yang dihasilkan disajikan pada Tabel 8.

Nama Parameter	Rentang Pencarian	Nilai Optimal Terpilih
num_leaves	20 – 50	37
max_depth	3 – 10	3
learning_rate	0.01 – 0.1	0.0930
scale_pos_weight	2.79 – 5.24	4.8136

ISSN 2598-9936 (online), https://ijins.umsida.ac.id, published by Universitas Muhammadiyah Sidoarjo

Indonesian Journal of Innovation Studies
Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

min_child_samples	10 – 30	15
lambda_l1 (L1)	0 – 5	1.6267
lambda_l2 (L2)	0 – 5	1.9434

Table 8. *Ruang Pencarian dan Hasil Parameter Terbaik*

2. Penerapan Balanced Strategy (Threshold Tuning)

Setelah model dilatih dengan parameter optimal, dilakukan evaluasi pada data uji (Test Set). Secara default, model klasifikasi menggunakan ambang batas (threshold) 0.5 untuk menentkan kelas. Namun, pada kasus medis dengan data ridak seimbang, ambang batas 0.5 seringkali tidak optimal karena cenderung menguntungkan kelas mayoritas.

Oleh karena itu, penelitian ini menerapkan strategi Threshold Tuning dengan mencari nilai ambang batas yang memaksimalkan skor F1-Score. Berdasarkan analisis kurva Precision-Recall, ditemukan bahwa threshold terbaik berada pada titik 0.7443 [18]. Dengan menggunakan ambang batas baru ini, prediksi model menjadi jauh lebih seimbang antara sensitivitas dan presisi. Evaluasi detail terhadap distribusi prediksi benar dan salah disajikan melalui Confusion Matrix pada Gambar 3.

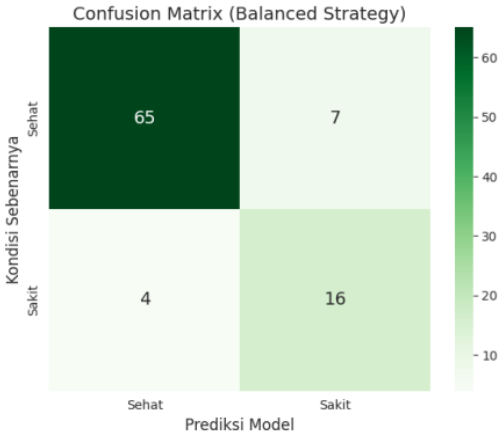


Figure 3. *Confusion Matrix pada Model Final*

Berdasarkan Gambar 3, model menunjukkan kemampuan yang sangat baik dalam mendeteksi kelas positif (Sakit). Dari total 20 sampel data subjek sakit yang ada di data uji, model berhasil mengklasifikasikan 16 subjek dengan benar (True Positive) dan hanya melewatkan 4 subjek (False Negative). Minimnya angka False Negative ini sangat krusial dalam konteks medis, karena membuktikan bahwa model memiliki sensitivitas yang tinggi untuk digunakan sebagai alat deteksi dini.

Selanjutnya, untuk mengukur kinerja model secara objektif pada data imbalanced, dilakukan analisis menggunakan Precision-Recall Curve yang ditampilkan pada Gambar 4.

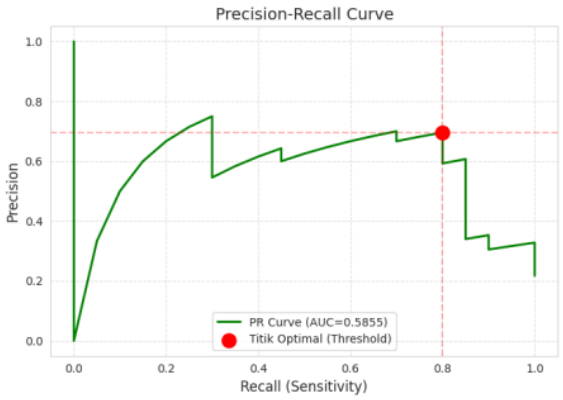


Figure 4. *Precision-Recall Curve dan Titik Optimal*

Gambar 4 memperlihatkan kurva PR dengan nilai PR-AUC (Area Under Curve) sebesar 0.5855. Titik merah pada grafik menunjukkan posisi threshold optimal (0.7443) yang dipilih model. Nilai PR-AUC 0.5855 ini tergolong baik mengingat

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

baseline (proporsi kelas positif) pada dataset ini hanya berkisar 0.22. Hal ini mengindikasikan bahwa kemampuan model dalam memisahkan kelas minoritas jauh melampaui tebakan acak (random chance), serta membuktikan efektivitas optimasi hyperparameter yang telah dilakukan.

3. Perbandingan Performa: Baseline vs. Final

Peningkatan kinerja yang diperoleh dari rangkaian proses seleksi fitur, optimasi hyperparameter, dan penyesuaian threshold dapat dilihat pada Tabel 9.

Metrik Evaluasi	Model Baseline	Model Final	Peningkatan
ROC-AUC	0.7418	0.8632	+0.1571
Akurasi	0.3823	0.8804	+13.86%
Recall	0.3228	0.8000	+47.72%
Presisi	0.3488	0.6957	+31.34%
F1-Score	0.7061	0.7442	+39.54%

Table 9. *Perbandingan Kinerja Akhir (Test Set)*

Tabel 9 menunjukkan lonjakan performa yang sangat signifikan. Skor ROC-AUC meningkat menjadi 0.8632, yang mengindikasikan kemampuan model yang sangat baik dalam membedakan antara subjek sehat dan subjek dengan gangguan mikrosirkulasi.

Peningkatan juga terlihat pada metrik Recall, yang melonjak dari 32,28% menjadi 80,00%. Hal ini membuktikan bahwa strategi pembobotan kelas scale_pos_weight sebesar 4.81 dikombinasikan dengan threshold tuning berhasil mengatasi masalah bias mayoritas yang sebelumnya dialami oleh model baseline. Model kini mampu mendeteksi 80% dari total kasus positif yang ada di data uji, menjadikan model ini layak digunakan sebagai alat bantu deteksi dini.

Peningkatan nilai Recall yang signifikan ini memiliki implikasi praktis yang sangat penting dalam konteks penerapan sistem deteksi dini berbasis wearable sensor. Recall yang tinggi menunjukkan bahwa sebagian besar subjek dengan kondisi sakit atau gangguan mikrosirkulasi berhasil teridentifikasi oleh sistem, sehingga risiko terlewatnya kasus positif (false negative) dapat ditekan secara substansial.

Dalam konteks medis, kesalahan false negative jauh lebih berbahaya dibandingkan false positif, karena dapat menyebabkan keterlambatan penanganan atau tidak terdeteksinya kondisi patalogis sejak tahap awal. Dengan kemampuan mendeteksi mayoritas kasus positif, model yang dikembangkan berpotensi digunakan sebagai alat skirining awal (early screening tool) dalam system pemantauan kesehatan berbasis wearable. System ini dapat berfungsi sebagai lapisan penyaring awal untuk mengidentidikasi individu yang memerlukan pemeriksaan lanjutan oleh tenaga medis, sehingga proses pengambilan keputusan klinis menjadi leboh efisien dan tepat sasaran, khususnya pada lingkungan dengan keterbatasan sumber daya keseharan.

Selain itu, peningkatan Recall yang disertai dengan nilai Presisi yang tetap berada pada tingkat yang memadai (69,57%) menunjukkan bahwa peningkatan sensitivitas model tidak dicapai dengan mengorbankan terlalu banyak kesalahan positif. Hal ini menjadikan sistem lebih dapat diterima secara praktis, karena peringatan yang dihasilkan memiliki tingkat keandalan yang cukup baik. Dengan demikian, model yang diusulkan tidak hanya unggul secara statistik, tetapi juga memiliki relevansi dan potensi implementasi nyata sebagai sistem pendukung deteksi dini gangguan mikrosirkulasi berbasis sensor wearable.

F. Analisis Feature Importance dan Interpretabilitas (SHAP)

Selain metrik evaluasi kinerja seperti akurasi dan ROC-AUC, aspek transparansi dalam pengambilan keputusan model memegang peranan vital, terutama untuk menghindari sifat black-box pada algoritma machine learning. Oleh karena itu, penelitian ini menerapkan metode SHAP (SHapley Additive exPlanations) guna menguraikan kontribusi marginal setiap fitur terhadap prediksi yang dihasilkan. Analisis ini bertujuan untuk memvalidasi apakah pola yang dipelajari oleh model selaras dengan logika medis atau fenomena data yang ada.

Gambaran mengenai fitur-fitur yang paling dominan dalam memengaruhi keputusan model disajikan melalui visualisasi SHAP Summary Plot pada Gambar 5. Grafik ini mengurutkan fitur dari posisi teratas hingga terbawah berdasarkan rata-rata

Vol. 27 No. 1 (2026): January
DOI: 10.21070/ijins.v27i1.1888

nilai absolut SHAP, yang mencerminkan tingkat kepentingan (importance) fitur tersebut. Sumbu horizontal menggambarkan dampak terhadap prediksi, di mana nilai positif mengindikasikan dorongan ke arah kelas Sakit, sedangkan nilai negatif mengarah pada kelas Sehat [19].

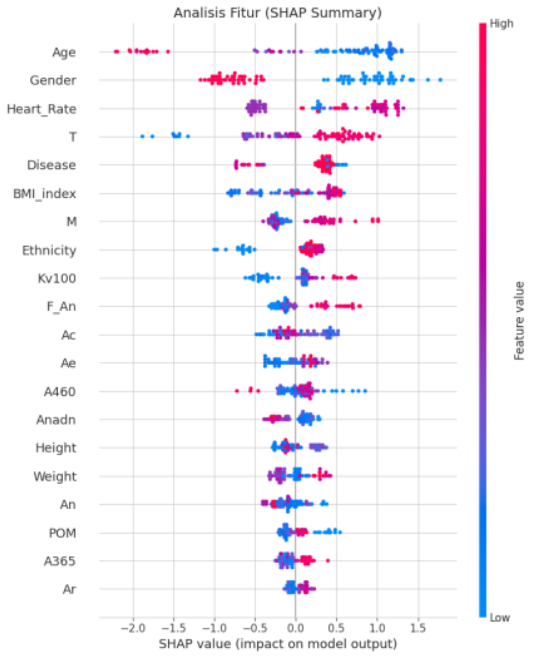


Figure 5. *SHAP Summary Plot*

Berdasarkan Gambar 5, terlihat bahwa variabel demografis mendominasi struktur keputusan model. Fitur Age (Usia) menempati peringkat pertama sebagai prediktor paling berpengaruh, diikuti oleh Gender di posisi kedua. Pola distribusi warna pada fitur Age memperlihatkan kecenderungan spesifik di mana titik-titik berwarna biru (merepresentasikan usia muda) lebih banyak tersebar di sisi kanan sumbu (nilai SHAP positif) [20]. Hal ini mengindikasikan bahwa pada dataset ini, subjek dengan usia yang lebih muda memiliki probabilitas yang lebih tinggi untuk terdeteksi sebagai kelas positif dibandingkan subjek yang lebih tua.

Selain faktor demografis, parameter fisiologis dan antropometri juga memberikan kontribusi signifikan. Pada fitur T (Suhu) dan BMI_index, model menangkap hubungan yang linier dan intuitif. Titik-titik berwarna merah yang merepresentasikan nilai suhu atau BMI tinggi cenderung berkumpul di sisi kanan grafik. Pola ini mengonfirmasi bahwa peningkatan suhu permukaan kulit dan indeks massa tubuh berkorelasi positif dengan peningkatan risiko deteksi kondisi patologis atau stres mikrosirkulasi. Sebaliknya, nilai yang rendah pada kedua fitur tersebut memberikan kontribusi negatif yang mendorong prediksi ke arah kondisi sehat. Kombinasi antara fitur demografis dan data sensor ini menunjukkan bahwa model mampu memanfaatkan ragam informasi secara komprehensif untuk membedakan karakteristik antar kelas.

Meskipun demikian, hasil yang diperoleh dalam penelitian ini masih membuka peluang pengembangan lanjutan. Salah satu arah pengembangan yang penting adalah pelaksanaan validasi klinis dengan melibatkan data dari populasi yang lebih luas dan beragam, serta pengujian langsung di lingkungan fasilitas kesehatan, guna memastikan bahwa performa dan pola keputusan model konsisten serta relevan secara medis. Selain itu, pendekatan yang diusulkan berpotensi untuk diintegrasikan secara real-time pada perangkat wearable, sehingga model tidak hanya berfungsi sebagai alat analisis offline, tetapi juga sebagai sistem pemantauan berkelanjutan (continuous monitoring) yang mampu memberikan peringatan dini terhadap indikasi gangguan mikrosirkulasi atau stres fisiologis. Integrasi tersebut diharapkan dapat meningkatkan nilai aplikatif penelitian ini dan mendukung implementasi sistem deteksi dini berbasis kecerdasan buatan dalam konteks kesehatan preventif dan klinis.

IV. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan mengenai optimasi hyperparameter algoritma LightGBM menggunakan Bayesian Optimization dan seleksi fitur untuk deteksi kondisi mikrosirkulasi berbasis data wearable LDF-FS, dapat ditarik simpulan sebagai berikut:

1. Algoritma LightGBM terbukti efektif dalam mengklasifikasikan kondisi mikrosirkulasi (Wellbeing vs Non-Wellbeing) pada dataset wearable yang memiliki karakteristik non-stationary dan imbalanced. Penerapan seleksi fitur berbasis feature importance (Total Gain)berhasil mereduksi dimensi data dari 34 atribut menjadi 22 fitur paling relevan, dengan fitur dominan meliputi faktor demografis (Age, BMI Index) dan parameter sensor (Temperature, Heart Rate).

2. Penerapan Bayesian Optimization secara signifikan meningkatkan kinerja model, khususnya dalam menangani ketidakseimbangan kelas. Penelitian nilai scale pos weight yang dinamis (sebesar 4.8136) melalui proses optimasi

mampu memaksa model untuk memberikan prioritas lebih tinggi pada kelas minoritas (Sakit), yang dibuktikan dengan lonjakan nilai Recall dari 32,28% (pada model baseline) menjadi 80,00% (pada model final)

3. Strategi Threshold Tuning yang diterapkan setelah optimasi hyperparameter berhasil menyeimbangkan nilai Precision dan Recall. Dengan menggeser ambang batas keputusan ke titik optimal 0.7443, model mencapai performa terbaik dengan nilai ROC-AUC sebesar 0.8632 dan Akurasi sebesar 88,04%. Peningkatan ini menunjukkan bahwa kombinasi pembobotan kelas dan penyesuaian threshold jauh lebih unggul dibandingkan penggunaan parameter default.

4. Analisis Interpretabilitas

   Analisis SHAP menunjukkan bahwa model mempelajari pola yang relevan secara klinis, dengan usia dan jenis kelamin sebagai predictor utama. Nilai SHAP mengindikasikan hubungan positif antara usia, suhu, jenis kelamin, sebagai predictor utama. Nilai SHAP mengindikasikan hubungan positif antara usia, suhu kulit, dan BMI terhadap risiko gangguan mikrosirkulasi, sehingga keputusan model selaras dengan logika medis dan bukan kebetulan statistic.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada seluruh pihak yang berkontribusi dalam penelitian ini, khususnya dalam penyediaan data wearable LDF-FS, bimbingan akademik, dan masukan ilmiha. Semoga hasil penelitian ini bermanfaat bagi pengembangan metode deteksi mikrosirkulasi yang akurat dan interpretative serta menjadi referensi bagi penelitian selanjutnya.

References

1. E. V. Zharkikh, Y. I. Loktionova, A. A. Fedorovich, A. Y. Gorshkov, dan A. V. Dunaev, “Assessment of blood microcirculation changes after COVID-19 using wearable laser doppler flowmetry,” *Diagnostics*, vol. 13, no. 5, p. 920, 2023, doi: 10.3390/diagnostics13050920.
2. L. Kralj dan H. Lenasi, “Wavelet analysis of laser doppler microcirculatory signals: Current applications and limitations,” *Frontiers in Physiology*, vol. 13, p. 1076445, 2023, doi: 10.3389/fphys.2023.1076445.
3. B. Mahesh, “Machine learning algorithms – a review,” *International Journal of Science and Research*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/SR20123143423.
4. M. N. Nguyen et al., “A wearable device dataset for mental health assessment using laser doppler flowmetry and fluorescence spectroscopy sensors,” *arXiv preprint*, arXiv:2502.00973, 2025, doi: 10.48550/arXiv.2502.00973.
5. M. A. Zöller dan M. F. Huber, “Benchmark and survey of automated machine learning frameworks,” *Journal of Artificial Intelligence Research*, vol. 70, pp. 409–472, 2021, doi: 10.1613/jair.1.12207.
6. J. Liu, Y. Wang, X. Zhang, dan Y. Chen, “A hybrid PSO-LightGBM model for disease prediction and feature selection,” *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–11, 2022, doi: 10.1155/2022/1287361.
7. E. Guldogan dan F. H. Yagin, “Interpretable machine learning for serum-based metabolomics in breast cancer diagnostics: Insights from multi-objective feature selection-driven LightGBM-SHAP models,” *Medicina*, vol. 61, no. 6, p. 1112, 2025, doi: 10.3390/medicina61061112.
8. Z. Noroozi, A. Orooji, dan L. Erfannia, “Analyzing the impact of feature selection methods on machine learning algorithms for heart disease prediction,” *Scientific Reports*, vol. 13, p. 22588, 2023, doi: 10.1038/s41598-023-49962-w.
9. A. Y. Aliefia dan M. I. Irawan, “Perbandingan metode Extreme Gradient Boosting (XGBoost) dan Light Gradient Boosting Machine (LightGBM) untuk mendeteksi fraud pada data klaim,” *Jurnal Sains dan Seni ITS*, 2025.
10. A. R. Raihan, A. W. Wardana, E. P. A. R., dan A. M. Rizki, “Perbandingan algoritma LightGBM dan ANN untuk menentukan kualitas anggur merah,” *JATI (Jurnal Mahasiswa Teknik)*, 2025.
11. H. K. Hendra et al., “Evaluasi performa Random Forest, XGBoost, dan LightGBM dalam diagnosis dini diabetes mellitus,” *Jurnal Penelitian Ilmu*, 2025. [Daring]. Tersedia: https://jurnal.polsri.ac.id/index.php/jupiter/article/view/10607
12. L. Hakim dan A. K. Zyen, “Optimasi model klasifikasi diabetes dengan stacking pada algoritma XGBoost dan LightGBM,” *JUPITER: Jurnal Penelitian Ilmu*, 2025. [Daring]. Tersedia: https://jurnal.polsri.ac.id/index.php/jupiter/article/view/10499
13. F. A. Wallad, “Perbandingan algoritma CatBoost dan LightGBM untuk diagnosis glioma: Studi kasus pada data TCGA,” *Repository AR-Raniry*, 2025. [Daring]. Tersedia: https://repository.ar-raniry.ac.id/id/eprint/46922/
14. A. N. Royana, Y. V. Via, dan C. A. Putra, “Evaluasi kinerja LightGBM dan CatBoost untuk prediksi churn berdasarkan dataset pelanggan layanan streaming musik,” *JATI (Jurnal Mahasiswa Teknik)*, 2025. [Daring]. Tersedia: https://www.ejournal.itn.ac.id/index.php/jati/article/download/14358/7954
15. E. E. Pardede, “Optimasi peramalan penjualan menu cafe menggunakan model hybrid Facebook Prophet dan LightGBM,” *Repository UPN Jatim*, 2025. [Daring]. Tersedia: https://repository.upnjatim.ac.id/43537/
16. A. P. Kasim et al., “Optimization of LightGBM model with Bayesian optimization for malware detection,” *Journal Unublitar*, 2025. [Daring]. Tersedia: http://journal.unublitar.ac.id/ilkomnika/index.php/ilkomnika/article/view/722
17. W. Hussain et al., “Ensemble genetic and CNN model-based image classification by enhancing hyperparameter tuning,” *Scientific Reports*, 2025.
18. R. Kochnev et al., “Optuna vs Code Llama: Are LLMs a new paradigm for hyperparameter tuning?” *arXiv preprint*, 2025. [Daring]. Tersedia: https://arxiv.org/abs/2504.06006
19. B. P. Lohani, A. Dagur, dan D. Shukla, “An efficient approach for diabetes classification using feature selection and hyperparameter tuning,” *Recent Advances in Electrical & Electronic Engineering*, 2025, doi: 10.2174/0123520965291885240315051751.
20. H. N. Fakhouri, S. Alawadi, F. M. Awaysheh, dan F. Hamad, “Novel hybrid success history intelligent optimizer with gaussian transformation: Application in CNN hyperparameter tuning,” *Cluster Computing*, 2024, doi: 10.1007/s10586-023-04161-0.