

APPLICATION OF INDEPENDENT COMPONENT ANALYSIS FOR FEATURE EXTRACTION IN HEART ARRHYTHMIA CLASSIFICATION USING LIGHT GRADIENT BOOSTING MACHINE

PENERAPAN INDEPENDENT COMPONENT ANALYSIS UNTUK EKSTRAKSI FITUR PADA KLASIFIKASI ARITMIA JANTUNG BERBASIS LIGHT GRADIENT BOOSTING MACHINE

Muhammad Rafif Danuja¹, Nada Ardelia^{2‡}, Mareska Radela Putra^{3‡}, Zanira Husna^{4‡}, Ayunda Rasyid Khairunnisa^{5‡}, Nida Khairunnissa^{6‡}, and Cici Suhaeni^{7‡}

^{1,2,3,4,5,6,7}Statistics and Data Study Program, IPB University, Indonesia

[‡]corresponding author: apip1505rafif@apps.ipb.ac.id

Copyright © 2026 Muhammad Rafif Danuja, Nada Ardelia, Mareska Radela Putra, Zanira Husna, Ayunda Rasyi Khairunnisa, Nida Khairunnissa, and Cici Suhaeni. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Cardiovascular disease, specifically arrhythmia, is a major health threat and one of the leading cause of death worldwide. Manually monitoring and detecting arrhythmias using an Electrocardiogram (ECG) is prone to human error due to the visual fatigue of medical staff. This research proposes a machine learning-based system to assist healthcare professionals in classifying arrhythmias more efficiently. The data used were obtained from the MIT-BIH Arrhythmia Database, focusing on the Modified Limb Lead II (MLII) channel. Data preprocessing was performed using cascading filters (High-pass, Low-pass, and Notch filters) and Z-score normalization to reduce noise, followed by beat segmentation based on three AAMI classes (N, S, and V). Principal Component Analysis (PCA) was applied prior to Independent Component Analysis (ICA) to reduce dimensionality. The primary focus of this research is the comparison of two ICA algorithms, namely FastICA and Picard, combined with morphological, statistical, and RR-interval features for ECG representation. To address class imbalance, sample weights were calculated from the training data and incorporated into the classification model. Classification was performed using Light Gradient Boosting Machine (LightGBM). The results show that the

proposed feature combination and LightGBM model achieved good performance for arrhythmia classification on real-world data, with FastICA achieved a macro F1-score of 0.7238, outperforming Picard which achieved 0.6963.

Keywords: Arrhythmia Classification, ECG, FastICA, LightGBM, Picard, SMOTE.

1. Pendahuluan

Gangguan irama jantung (aritmia) sebagai penyakit kardiovaskular merupakan salah satu tantangan kesehatan global yang paling mematikan. Keterlambatan dalam mengidentifikasi kondisi ini secara dini dapat memicu komplikasi yang fatal, salah satunya henti jantung mendadak. Meski analisis elektrokardiogram (EKG) merupakan prosedur standar untuk mendeteksi secara klinis, masih terdapat kendala dalam interpretasi manual oleh tenaga medis akibat kompleksitas sinyal dan risiko kesalahan manusia akibat kelelahan (Hannun et al., 2019). Hal ini mendorong perlunya pengembangan sistem klasifikasi otomatis berbasis *machine learning* sebagai solusi diagnostik yang lebih akurat dan efisien.

MIT-BIH *Arrhythmia Database* banyak digunakan sebagai standar global dalam penelitian klasifikasi aritmia. Namun, sinyal EKG pada dataset tersebut masih mengandung berbagai *noise* sehingga diperlukan tahapan *preprocessing* sebelum dilakukan ekstraksi fitur (Malghan & Hota, 2020). Salah satu metode ekstraksi fitur yang banyak digunakan adalah *Independent Component Analysis* (ICA), yaitu metode dekomposisi sinyal untuk memisahkan komponen independen pada sinyal observasi. Selain fitur hasil ICA, penelitian sebelumnya menunjukkan bahwa fitur statistik, morfologi, dan *RR-interval* juga memiliki kemampuan diskriminatif dalam membedakan jenis aritmia (Sahoo et al., 2020; Rahul et al., & Arumuganathan, 2021). Selain itu, klasifikasi aritmia juga menghadapi ketidakseimbangan distribusi kelas karena kelas normal mendominasi jumlah data dibandingkan dengan kelas lainnya. Kondisi ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas sehingga diperlukan strategi penanganan ketidakseimbangan kelas yang tepat agar performa klasifikasi pada setiap kelas dapat dievaluasi secara akurat.

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan sistem klasifikasi aritmia berbasis gabungan fitur *Independent Component Analysis* (ICA), fitur morfologi, statistik deskriptif, dan *RR-interval* pada sinyal EKG. Penelitian ini berfokus pada klasifikasi tiga jenis kelas aritmia AAMI, yaitu Normal (N), Supraventricular (S), dan Ventricular (V), juga membandingkan algoritma FastICA dan Picard dalam menghasilkan representasi fitur yang optimal. Setelah itu, fitur diklasifikasikan

menggunakan *Light Gradient Boosting Machine* (LightGBM) dengan penanganan ketidakseimbangan kelas melalui penerapan *sample weights* berdasarkan distribusi kelas pada data latih. Evaluasi dilakukan menggunakan *macro F1-score* pada skema *inter-patient split* untuk membandingkan performa klasifikasi kedua pendekatan ICS pada kata EKG nyata.

2. Metodologi

2.1 Bahan dan Data

Penelitian ini menggunakan data sekunder berupa sinyal elektrokardiogram (EKG) yang bersumber dari MIT-BIH *Arrhythmia Database* (Moody & Mark, 2001), diakses melalui portal PhysioNet. Dataset ini memuat 48 pasien dengan durasi 30 menit per rekaman pada frekuensi sampling 360 Hz. Sebanyak empat rekaman dengan alat pacu jantung (*record* 102, 104, 107, dan 217) dikeluarkan dari analisis karena morfologi sinyalnya bersifat artifisial dan tidak merepresentasikan ritme jantung alami (de Chazal et al., 2004), sehingga total rekaman yang dianalisis berjumlah 44 rekaman.

2.2 Prosedur Penelitian

Analisis data dilakukan menggunakan bahasa pemrograman Python. Tahapan analisis data yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Praproses dan eksplorasi data.

Praproses sinyal dilakukan melalui dua tahapan utama, yaitu *filtering* dan *normalisasi*, yang efektivitasnya divalidasi melalui visualisasi perbandingan sinyal sebelum dan sesudah praproses serta analisis *Power Spectral Density* (PSD).

- a. *Filtering*, yaitu proses reduksi noise dilakukan menggunakan metode *cascading filters* (filter bertingkat) yang dirancang untuk mengatasi tiga jenis gangguan utama pada EKG. Filter ini mencakup
 - 1) *High-pass filter* dengan *cutoff* 0.5 Hz untuk mengeliminasi *baseline wander* yang umumnya disebabkan oleh pergerakan respirasi pasien.
 - 2) *Low-pass filter* dengan *cutoff* 45 Hz untuk meredam *noise* berfrekuensi tinggi seperti kontraksi otot (Elektromiogram/EMG).
 - 3) *Notch filter* pada frekuensi 60 Hz untuk menghilangkan gangguan yang disebabkan oleh arus listrik alat perekam.
- b. *Normalisasi* dengan *Z-score* agar setiap rekaman memiliki skala yang seragam sesuai Persamaan 1

$$Z = \frac{x - \mu}{\sigma}. \quad (1)$$

dengan z adalah sinyal EKG yang telah dinormalisasi (memiliki rata-rata 0 dan standar deviasi 1). x adalah nilai pengamatan asli (amplitudo sinyal EKG mentah pada waktu tertentu). μ adalah rata-rata dari seluruh nilai pengamatan. σ adalah standar deviasi yang merepresentasikan persebaran data.

2. Segmentasi *Beat* dan Pemetaan Kelas AAMI

- a. Segmentasi *Beat*: Setiap *beat* disegmentasi menggunakan jendela tetap 150 ms sebelum dan 450 ms sesudah puncak R dari file anotasi, menghasilkan segmen sepanjang 216 sampel per *beat* pada frekuensi sampling 360 Hz. Pendekatan asimetris ini dipilih agar seluruh komponen gelombang P, QRS, dan T tercakup dalam satu segmen, sekaligus memungkinkan ekstraksi *pre-RR* dan *post-RR* interval sebagai fitur kontekstual ritme jantung (Dias et al., 2021).
- b. Pemetaan Standar AAMI: Anotasi *beat* asli dipetakan ke standar AAMI (*Association for the Advancement of Medical Instrumentation*) menjadi 5 kelas: N (Normal), S (Supraventricular), V (Ventricular), F (Fusion), dan Q (Unknown), untuk menyamakan *ground truth* dengan standar internasional dalam klasifikasi aritmia (Irid et.al, 2024). Penelitian ini memodelkan tiga kelas, yaitu N, S, dan V, dengan mengecualikan kelas F dan Q. Kelas F dikeluarkan karena merepresentasikan *beat* hibrida dengan morfologi intermediet antara *beat* N dan V yang ambigu secara diagnostik (Luz et al., 2016). Sementara itu, kelas Q dieksklusi karena secara klinis mencakup *unclassifiable beat* dan *paced beat* yang dihasilkan alat pacu jantung buatan dengan morfologi unik yang tidak mencerminkan aritmia alami, sehingga berisiko memperkenalkan bias yang tidak relevan secara klinis (de Chazal et al., 2004).

3. Pembagian Data

Rekaman dibagi menggunakan skema inter-patient split yang pertama kali diusulkan oleh de Chazal et al. (2004) menjadi DS1 sebagai data latih dan DS2 sebagai data uji, masing-masing terdiri dari 22 rekaman. Pendekatan ini bertujuan untuk memastikan tidak ada data dari pasien yang sama yang muncul pada kedua subset sekaligus, sehingga model yang dihasilkan lebih mampu menggeneralisasi pola aritmia lintas pasien.

Tabel 1: Skema *interpatient-split*.

Dataset	Rekaman
DS1 (<i>Training</i>)	'101','106','108','109','112','114','115', '116','118','119','122','124','201','203', '205','207','208','209','215','220','223', '230'
DS2 (<i>Testing</i>)	'100','103','105','111','113','117','121', '123','200','202','210','212','213','214', '219','221','222','228','231','232','233', '234'

4. Ekstraksi Fitur *Handcrafted*

Dari setiap segmen beat, diekstraksi tiga kelompok fitur:

- Fitur morfologi (7 fitur): amplitudo R, amplitudo S, amplitudo T, energi QRS, luas QRS, lebar QRS, dan rasio RT.
- Fitur statistik (20 fitur): mean, standar deviasi, varians, skewness, kurtosis, persentil (P10, P25, P75, P90), *interquartile range* (IQR), *root mean square* (RMS), *peak-to-peak amplitude*, *zero crossing rate*, parameter Hjorth (*activity*, *mobility*, *complexity*), *spectral entropy*, *dominant frequency*, *spectral centroid*, dan rasio frekuensi LF/HF.
- Fitur interval RR (2 fitur): RR sebelumnya dan RR berikutnya yang dinormalisasi terhadap rata-rata RR dataset latih (DS1).

5. Reduksi Dimensi: PCA dan ICA

Waveform 216 sampel terlebih dahulu dikompres menggunakan PCA *whitening* dengan retensi varians sebesar 95%. Komponen PCA yang dihasilkan kemudian diproses oleh *Independent Component Analysis* (ICA) untuk memisahkan sumber sinyal yang statistik independen. Penelitian ini membandingkan dua implementasi algoritma ICA:

- FastICA (*sklearn*): menggunakan fungsi kontras *logcosh* dengan maksimum iterasi 2.000, dan toleransi konvergensi $1e-4$ (Hyvärinen & Oja, 2000; Pedregosa et al., 2011).
- Picard: menggunakan implementasi *python-picard* dengan mode non-orthogonal, maksimum iterasi 500, dan toleransi $1e-7$ (Ablin et al., 2018).

Setelah itu, fitur ICA digabungkan dengan fitur *handcrafted* menjadi satu matriks fitur terpadu yang digunakan untuk klasifikasi. Seluruh fitur distandarisasi kembali menggunakan *StandardScaler* sebelum masuk ke proses klasifikasi.

6. Penanganan *Imbalanced Data*

Distribusi kelas pada dataset aritmia bersifat tidak seimbang (*imbalanced*), di mana kelas N memiliki jumlah sampel yang jauh lebih besar dibandingkan kelas S dan V. Untuk menangani hal tersebut diterapkan strategi *class weight* yang dihitung langsung dari distribusi kelas pada data latih (DS1) menggunakan skema *balanced weighting* sesuai Persamaan 2.

$$w_c = \frac{N}{K \cdot n_c} \quad (2)$$

di mana w_c adalah bobot untuk kelas c , N adalah total jumlah sampel pada data latih, K adalah jumlah kelas yang dimodelkan, dan n_c adalah jumlah sampel pada kelas c . Bobot yang dihasilkan kemudian digunakan sebagai *class weight* dalam fungsi *loss* selama pelatihan, sehingga setiap sampel dari kelas minoritas berkontribusi lebih besar terhadap pembaruan gradien dibandingkan sampel dari kelas mayoritas.

6. Pemodelan Klasifikasi dengan LightGBM dan *Hyperparameter Tuning*

Pada tahap pemodelan, penelitian ini menggunakan algoritma LightGBM melalui *library lightgbm* untuk melakukan klasifikasi jenis aritmia berdasarkan

fitur hasil ekstraksi sinyal ECG. LightGBM merupakan algoritma *ensemble* berbasis *gradient boosting* yang membangun model prediktif secara bertahap melalui serangkaian pohon keputusan dengan *pendekatan leaf-wise tree growth* dan *histogram-based splitting*, sehingga mampu meningkatkan efisiensi komputasi sekaligus mempertahankan performa prediksi yang tinggi (Ke et al., 2017). Untuk memperoleh konfigurasi optimal, dilakukan hyperparameter tuning menggunakan Bayesian Optimization melalui framework Optuna, di mana data pelatihan dibagi kembali menjadi subset training dan validation untuk mengevaluasi setiap kombinasi parameter. Ruang pencarian *hyperparameter* yang digunakan ditunjukkan pada Tabel 2, dan kombinasi terbaik dipilih berdasarkan performa pada data validation untuk membangun model klasifikasi akhir.

Tabel 2: Ruang pencarian *hyperparameter*.

Nama	Nilai
<i>n_estimators</i>	200 – 1000
<i>Learning rate</i>	0,01 – 0,20 (log scale)
<i>max_depth</i>	{-1, 6, 8, 10, 12}
<i>num_leaves</i>	15 – 255
<i>min_child_samples</i>	5 – 100
<i>subsamples</i>	0,50 – 1,00
<i>colsample_bytree</i>	0,50 – 1,00
<i>reg_alpha</i>	10^{-4} – 10 (log scale)
<i>reg_lambda</i>	10^{-4} – 10 (log scale)

7. Evaluasi Model

Kinerja model dinilai berdasarkan *Confusion Matrix* menggunakan beberapa metrik klasifikasi untuk mengukur akurasi dan kemampuan deteksi tiap kelas (Putra et al., 2022).

- a. *Accuracy*: menunjukkan proporsi prediksi yang benar terhadap seluruh data uji.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%. \quad (3)$$

- b. *Precision*: mengukur ketepatan model dalam memprediksi kelas positif.

$$Precision = \frac{TP}{TP + FP} \times 100. \quad (4)$$

- c. *Recall*: menunjukkan kemampuan model dalam mendeteksi seluruh sampel positif.

$$Recall = \frac{TP}{TP + FN} \times 100. \quad (5)$$

- d. *F1-Score*: merupakan rata-rata harmonik antara *precision* dan *recall* (Pambudi et al., 2025).

$$F1_{score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \quad (6)$$

Evaluasi model dilakukan pada dataset DS2 sebagai data uji yang terpisah dari data pelatihan untuk menilai kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya. Selain performa klasifikasi, perbandingan antara FastICA dan Picard juga dilakukan berdasarkan kualitas dekomposisi ICA melalui *mean absolute excess kurtosis* (non-Gaussianity), serta efisiensi komputasi yang diukur dari waktu *fitting* ICA dan total waktu eksekusi *pipeline*.

8. McNemar's Test

Perbedaan performa antara model FastICA dan Picard diuji secara statistik menggunakan McNemar's Test (McNemar, 1947; Dietterich, 1998). Uji ini bekerja dengan membentuk tabel kontingensi 2×2 berdasarkan pasangan prediksi discordant antar model, dengan statistik uji:

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}. \quad (7)$$

di mana b dan c masing-masing adalah jumlah sampel yang hanya diprediksi benar oleh satu model. Jika jumlah pasangan *discordant* ($b + c$) < 25, digunakan uji eksak berbasis distribusi binomial sebagai pengganti. Hipotesis nol yang diuji adalah $H_0: b = c$, dengan tingkat signifikansi $\alpha = 0,05$. Uji dilakukan secara keseluruhan maupun per kelas (*one-vs-rest*).

9. Analisis Error

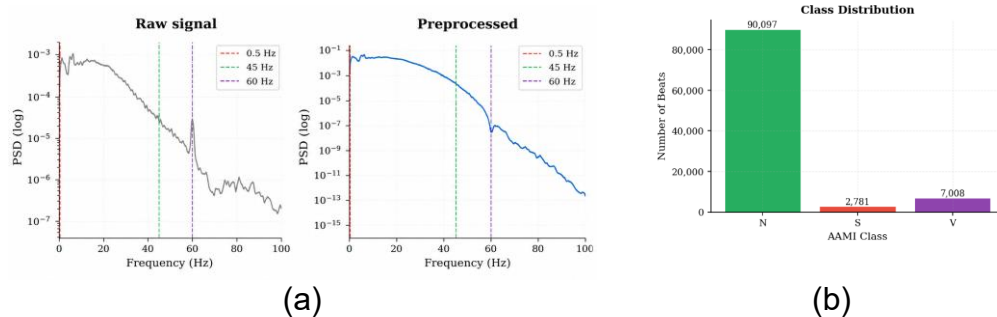
Analisis error pada kelas S dilakukan melalui proyeksi ruang fitur FastICA pada DS2 menggunakan *t-Distributed Stochastic Neighbor Embedding* (t-SNE), yaitu metode yang memetakan data berdimensi tinggi ke ruang dua dimensi dengan mempertahankan kemiripan antar sampel (van der Maaten & Hinton, 2008). Proyeksi dilakukan dengan *perplexity* = 40 dan 1.000 iterasi pada subsampel acak maksimum 3.000 sampel, kemudian divisualisasikan berdasarkan label *ground truth* dan label prediksi untuk mengidentifikasi pola *overlap* antar kelas yang berpotensi menjadi sumber misklasifikasi.

3. Hasil dan Pembahasan

3.1 Eksplorasi dan Praproses Data

Pada tahap awal eksplorasi, analisis morfologi sinyal EKG dilakukan pada cuplikan record 100 (0–10 detik) untuk mengidentifikasi komponen fisiologis dan gangguan fisis sinyal. Sinyal mentah menunjukkan ketidakstabilan baseline pada rentang −0,25 hingga −0,50 mV, mengindikasikan adanya baseline wander dan noise frekuensi rendah yang berpotensi mengganggu ekstraksi fitur. Untuk mengatasi hal tersebut, sinyal diproses menggunakan *cascading filters* (HPF 0,5 Hz, LPF 45 Hz, dan Notch 60 Hz) serta normalisasi

Z-score, yang berhasil menstabilkan *baseline* di sekitar nilai 0 dan menghasilkan morfologi gelombang P, QRS, dan T yang jauh lebih bersih dan distingtif. Efektivitas *pipeline* praproses tersebut serta komposisi distribusi kelas dataset disajikan pada Gambar 1.



Gambar 1: Plot *Power Spectral Density* (PSD) sebelum dan sesudah praproses (a), plot jumlah beat per kelas (b).

Validasi efektivitas *filtering* dikonfirmasi melalui analisis *Power Spectral Density* (PSD) metode Welch pada Gambar 1(a), di mana sinyal pasca-praproses menunjukkan eliminasi signifikan pada interferensi 60 Hz, noise EMG di atas 45 Hz, dan *baseline wander* di bawah 0,5 Hz hingga turun beberapa orde magnitudo, sehingga energi yang tersisa murni merepresentasikan informasi fisiologis jantung pada rentang 0,5–45 Hz. Sementara itu, Gambar 1(b) mengonfirmasi ketidakseimbangan kelas yang ekstrem, dengan kelas N mendominasi sebesar 90.097 beat dibandingkan kelas V (7.008 beat) dan S (2.781 beat), yang memperkuat keputusan untuk hanya mempertahankan ketiga kelas tersebut dalam pemodelan. Selanjutnya, dataset dibagi menggunakan skema inter-patient split menjadi DS1 (data latih) dan DS2 (data uji) dengan proporsi kelas yang representatif, sehingga DS2 dapat berfungsi sebagai data uji yang valid dan bebas bias.

3.2 *Independent Component Analysis* (ICA) dengan FastICA dan Picard

3.2.1 Ekstraksi Fitur Menggunakan FastICA dan Picard

Ekstraksi fitur diawali dengan reduksi dimensi menggunakan PCA pada seluruh segmen beat ECG hasil praproses, guna mengurangi redundansi antarfitur, menekan noise residual, serta meningkatkan stabilitas dan efisiensi dekomposisi ICA. Dari proses reduksi tersebut, sebanyak 14 komponen utama terpilih mampu merepresentasikan 95,07% variasi data asli, yang kemudian digunakan sebagai masukan pada proses dekomposisi ICA menggunakan metode FastICA dan

Picard. Ringkasan hasil ekstraksi fitur dari kedua metode disajikan pada Tabel 3.

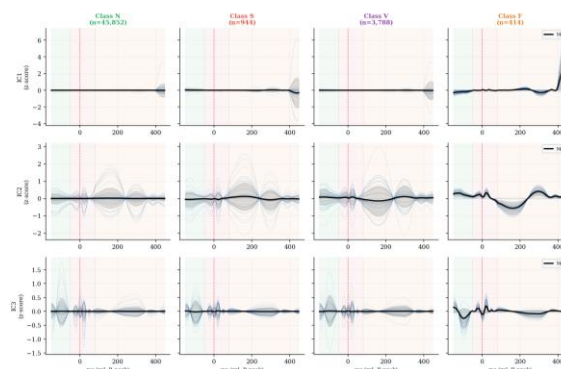
Tabel 3: Daftar fitur hasil ekstraksi sinyal ECG.

Metode	Variansi PCA	Jumlah Komponen	Excess Kurtosis	Waktu ICA	Total Waktu
FastICA	95.07%	14	36.565	0.1 s	0.4 s
Picard	95.07%	14	35.992	5.2 s	5.4 s

Berdasarkan Tabel 3, kedua metode menghasilkan jumlah komponen yang sama karena menggunakan hasil reduksi PCA yang identik. Perbedaan utama terlihat pada kualitas pemisahan sinyal dan efisiensi komputasi. FastICA menghasilkan nilai excess kurtosis yang sedikit lebih tinggi (36,565) dibandingkan Picard (35,992), yang menunjukkan tingkat non-Gaussianitas komponen yang lebih baik. Selain itu, FastICA membutuhkan total waktu pemrosesan hanya 0,4 detik, jauh lebih cepat dibandingkan Picard yang memerlukan 5,4 detik. Hasil ini menunjukkan bahwa FastICA mampu menghasilkan fitur yang setara bahkan sedikit lebih baik, dengan biaya komputasi yang jauh lebih rendah.

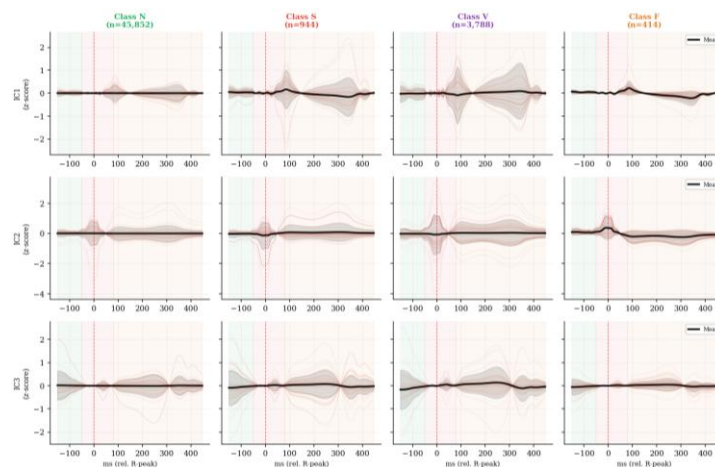
3.2.2 Dekomposisi Fitur FastICA dan Picard

Proses dekomposisi dilakukan menggunakan ICA, karakteristik fitur independen selanjutnya dianalisis melalui visualisasi bentuk gelombang (*waveforms*) pada tiga komponen utama pertama, yaitu IC1, IC2, dan IC3 untuk setiap kelas detak jantung AAMI. Visualisasi tersebut disajikan pada Gambar 4 untuk FastICA dan Gambar 5 untuk Picard.



Gambar 4: Visualisasi *reconstructed ICA waveforms* hasil FastICA pada tiga komponen independen utama (IC1–IC3) untuk setiap kelas AAMI.

Pada hasil dekomposisi menggunakan FastICA, komponen independen yang terbentuk menunjukkan pola pemisahan fitur yang cukup jelas antarkelas detak jantung. Komponen IC1 terlihat mampu mengisolasi pola sinyal tertentu pada kelas F (*Fusion beat*), yang ditunjukkan dengan adanya deviasi amplitudo yang lebih menonjol dibandingkan kelas lainnya. Sementara itu, IC2 dan IC3 menangkap pola osilasi fisiologis pada rentang setelah puncak R yang berkaitan dengan aktivitas repolarisasi ventrikel. Variasi antargelombang pada kelas N, S, dan V masih mempertahankan pola rata-rata yang relatif stabil di sekitar garis dasar, sedangkan kelas F menunjukkan bentuk morfologi yang lebih berbeda sejak awal segmen sinyal.



Gambar 5: Visualisasi *reconstructed ICA waveforms* hasil Picard ICA pada tiga komponen independen utama (IC1–IC3) untuk setiap kelas AAMI.

Berbeda dengan FastICA, hasil dekomposisi menggunakan Picard menunjukkan pola gelombang yang cenderung lebih halus dan homogen antarkelas. Variasi amplitudo pada setiap komponen tampak lebih kecil dengan rentang nilai yang lebih sempit. Selain itu, pola rata-rata pada IC1, IC2, dan IC3 terlihat lebih landai sehingga perbedaan karakteristik antarkelas tidak terlihat se jelas pada FastICA. Kondisi ini mengindikasikan bahwa sebagian detail informasi sinyal mengalami reduksi selama proses optimisasi Picard, sehingga struktur morfologi khas pada beberapa kelas aritmia menjadi kurang menonjol.

Fitur hasil ICA kemudian digabungkan kembali dengan fitur lainnya sehingga menghasilkan representasi fitur akhir sebanyak 43 fitur (14 komponen ICA, 7 morfologi, 2 RR-interval, 20 statistik). Dataset *training* memiliki ukuran 50.998 sampel dengan 43 fitur, sedangkan dataset testing memiliki ukuran 49.690 sampel dengan 43 fitur. Visualisasi fitur ICA pada IC1 dan IC2 menunjukkan bahwa beberapa

kelas mulai membentuk pola pemisahan setelah transformasi ICA dilakukan.

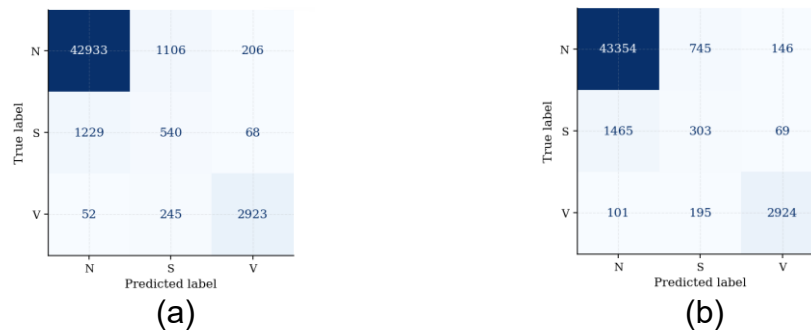
3.3 Pemodelan Klasifikasi dengan LightGBM

Model klasifikasi pada tahap ini dibangun menggunakan algoritma LightGBM dengan dua representasi fitur yang berbeda, yaitu fitur hasil FastICA dan Picard. Proses pelatihan dilakukan menggunakan hyperparameter terbaik hasil optimasi Optuna serta mekanisme threshold tuning pada data pelatihan. Pada pengujian menggunakan dataset DS2, metode FastICA–LightGBM menghasilkan *macro F1-score* sebesar 0,7238, sedangkan metode Picard–LightGBM menghasilkan *macro F1-score* sebesar 0,6963. Rincian performa klasifikasi untuk setiap kelas disajikan pada Tabel 4.

Tabel 4: Evaluasi model LGBM pada DS2.

Kelas	Recall		Precision		F1-score	
	FastICA	Picard	FastICA	Picard	FastICA	Picard
N	0,9703	0,9799	0,9710	0,9651	0,9707	0,9724
S	0,2940	0,1649	0,2856	0,2438	0,2897	0,1968
V	0,9078	0,9081	0,9143	0,9315	0,9110	0,9196

Berdasarkan Tabel 4, kedua model menunjukkan pola performa yang relatif serupa. Kelas N dan V memiliki tingkat klasifikasi yang sangat baik pada kedua metode, dengan nilai F1-score di atas 0,91. Pada kelas N, metode Picard–LightGBM memperoleh F1-score sedikit lebih tinggi (0,9724) dibandingkan FastICA–LightGBM (0,9707), sedangkan pada kelas V, Picard–LightGBM juga menghasilkan F1-score yang lebih tinggi sebesar 0,9196 dibandingkan 0,9110 pada FastICA–LightGBM. Pada kelas S yang merupakan kelas minoritas, FastICA–LightGBM menunjukkan kinerja yang lebih baik dengan *recall* sebesar 0,2940 dan F1-score sebesar 0,2897, sedangkan Picard–LightGBM hanya mencapai *recall* sebesar 0,1649 dan F1-score sebesar 0,1968. Peningkatan kemampuan mendeteksi kelas S tersebut menyebabkan FastICA–LightGBM memperoleh *macro F1-score* yang lebih tinggi (0,7238) dibandingkan Picard–LightGBM (0,6963). Hasil ini menunjukkan bahwa representasi fitur FastICA lebih efektif dalam membantu LightGBM mengenali pola pada kelas yang sulit diklasifikasikan, khususnya *heartbeat supraventricular* (S).



Gambar 6: *Confusion matrix* (a) FastICA-LightGBM dan (b) Picard-LightGBM.

Hasil *confusion matrix* pada Gambar 6 memperkuat pola performa yang ditampilkan pada Tabel 3. Pada kedua model, sebagian besar data kelas N berhasil diklasifikasikan dengan benar, yaitu sebanyak 42.933 *beat* pada FastICA dan 43.354 *beat* pada Picard. Kelas V juga menunjukkan jumlah prediksi benar yang tinggi, yaitu 2.923 *beat* pada FastICA dan 2.924 *beat* pada Picard. Hal ini menunjukkan bahwa kedua model mampu mengenali pola *heartbeat normal* dan *ventricular* dengan cukup baik. Kemiripan prediksi *beat* terlihat pada kelas S yang menunjukkan sebanyak 1.229 *beat* pada FastICA dan 1.465 *beat* pada Picard diklasifikasikan sebagai kelas N. Kondisi ini menunjukkan bahwa kelas S masih memiliki *overlap* karakteristik dengan kelas normal sehingga model cenderung mengarahkannya ke kelas mayoritas.

Hasil penelitian ini dapat dibandingkan dengan beberapa penelitian terdahulu yang menggunakan skema *inter-patient split* pada MIT-BIH *Arrhythmia Database*. De Chazal et al. (2004) menggunakan fitur morfologi dan RR-interval dan melaporkan sensitivitas sebesar 75,9% untuk kelas *supraventricular* (S) dan 77,7% untuk kelas *ventricular* (V). Penelitian yang lebih baru oleh Zhou et al. (2023) melaporkan akurasi keseluruhan sebesar 99,88%, sensitivitas 97,66%, dan *positive predictive value* 96,14% pada kelas S menggunakan model *deep learning* berbasis representasi temporal ECG. Dibandingkan dengan penelitian-penelitian tersebut, penelitian ini tidak berfokus pada pengembangan arsitektur klasifikasi baru, melainkan pada evaluasi kualitas representasi fitur melalui perbandingan FastICA dan Picard yang dikombinasikan dengan fitur morfologi, statistik, dan RR-interval sebelum diklasifikasikan menggunakan LightGBM.

3.4 McNemar's Test

Hasil pengujian menunjukkan nilai statistik sebesar 42,06 dengan $p\text{-value} < 0,001$, yang mengindikasikan adanya perbedaan yang signifikan secara statistik antara kedua metode. Dari tabel kontingensi, diperoleh $b=310$ (FastICA benar, Picard salah) dan $c=495$ (Picard benar, FastICA salah), yang menunjukkan bahwa Picard lebih sering menghasilkan prediksi benar dibanding FastICA pada sampel-sampel yang berbeda keputusan.

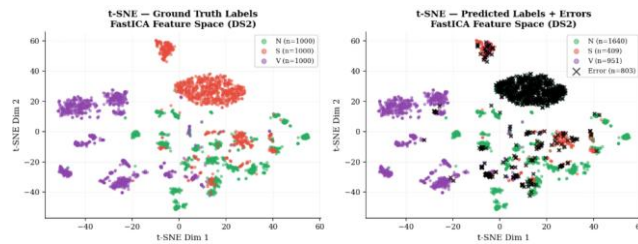
FastICA Benar	a: Keduanya Benar 46,086	b: FI Benar PC Salah 310
FastICA Salah	c: FI Salah PC Benar 495	d: Keduanya Salah 2,411
	Picard Benar	Picard Salah

Gambar 7: Hasil McNemar's *test*.

Namun, hasil ini perlu mempertimbangkan karakteristik data yang tidak seimbang (*imbalanced*). McNemar's test secara implisit lebih merefleksikan kontribusi kelas mayoritas terhadap akurasi global, sehingga keunggulan Picard pada uji ini lebih mencerminkan performa berbasis *accuracy* daripada keseimbangan antar kelas. Sebaliknya, evaluasi menggunakan *macro-F1* menunjukkan bahwa FastICA memiliki performa yang lebih baik karena memberikan kontribusi yang lebih seimbang pada seluruh kelas, termasuk kelas minoritas. Dengan demikian, meskipun Picard unggul secara signifikan pada uji McNemar, FastICA lebih unggul dalam metrik *macro-F1* yang lebih representatif untuk kondisi data *imbalanced*.

3.5 Analisis Error

Sebagai pelengkap evaluasi statistik, analisis error dilakukan secara visual menggunakan proyeksi t-SNE pada ruang fitur FastICA untuk DS2. Panel kiri pada Gambar 9 bagian (a) menampilkan distribusi ground truth label, di mana ketiga kelas yaitu N ($n=1.000$), S ($n=1.000$), dan V ($n=1.000$) membentuk kluster yang relatif terpisah meskipun terdapat tumpang tindih pada beberapa wilayah batas antarkluster, khususnya antara kelas N dan S.



Gambar 9: Proyeksi t-SNE ruang fitur FastICA pada DS2: distribusi *ground truth* label (a), distribusi prediksi model beserta sebaran error (b).

Gambar 9 bagian (b) menampilkan prediksi model beserta sebaran error ($n=803$), yang terkonsentrasi di wilayah perbatasan antara kelas N dan S, mengonfirmasi bahwa kesalahan klasifikasi utamanya terjadi akibat *overlapping* distribusi fitur pada kedua kelas tersebut. Kelas V menunjukkan separabilitas yang lebih baik dengan error yang relatif sedikit dan tersebar, sementara kelas S menjadi sumber error terbesar, konsisten dengan performa *macro-F1* yang lebih rendah pada kelas minoritas. Pola ini memperkuat temuan McNemar's test bahwa perbedaan performa antar metode berpusat pada kemampuan masing-masing dalam menangani ambiguitas di wilayah keputusan kelas N dan S, bukan pada kelas V yang secara geometris lebih mudah dipisahkan.

4. Simpulan dan Saran

Penelitian ini terbukti mampu membangun sistem klasifikasi aritmia berbasis sinyal elektrokardiogram (EKG) dengan kombinasi fitur *Independent Component Analysis* (ICA), fitur morfologi, statistik deskriptif, dan *RR-interval*. *Cascading filter* dan normalisasi *Z-score* digunakan pada tahapan *preprocessing* yang menghasilkan representasi sinyal yang lebih stabil untuk proses ekstraksi fitur. Selain itu, penanganan ketidakseimbangan kelas dilakukan melalui penerapan *sample weights* berdasarkan distribusi kelas pada data latih sehingga proses pembelajaran model menjadi lebih seimbang.

Hasil penelitian menunjukkan bahwa model LightGBM mampu menghasilkan performa klasifikasi aritmia yang baik pada data EKG nyata. Perbandingan antara dua algoritma ICA berdasarkan nilai *macro F1-score* menunjukkan bahwa FastICA memberikan performa yang sedikit lebih baik dibandingkan dengan Picard dengan nilai masing-masing sebesar 0,7238 dan 0,6963. Model mampu mengklasifikasikan kelas Normal(N), *supraventricular* (S), dan *ventricular* (V) dengan cukup baik berdasarkan evaluasi *macro F1-score*. Dapat disimpulkan bahwa kombinasi fitur ICA, morfologi, statistik deskriptif, dan *RR-interval* terbukti efektif dalam merepresentasikan karakteristik sinyal EKG untuk klasifikasi aritmia.

Daftar Pustaka

- Ablin, P., Cardoso, J.-F., & Gramfort, A. (2018). Faster independent component analysis by preconditioning with Hessian approximations. *IEEE Transactions on Signal Processing*, 66(15), 4040–4049. <https://doi.org/10.1109/TSP.2018.2844203>.
- De Chazal, P., O'Dwyer, M., & Reilly, R. B. (2004). Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*, 51(7), 1196–1206. <https://doi.org/10.1109/TBME.2004.827359>.
- Dias, F. M., Monteiro, H. L. M., Cabral, T. W., Naji, R., Kuehni, M., Luz, E. J. S. (2021). Arrhythmia classification from single-lead ECG signals using the inter-patient paradigm. *Computer Methods and Programs in Biomedicine*. 202: 105948. <https://doi.org/10.1016/j.cmpb.2021.105948>.
- Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10(7), 1895–1923.
- Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, 25(1), 65-69.
- Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5), 411-430.
- Irid, S. M. H., Moussaoui, D., Hadjila, M. and Azzoug, O. (2024). Classification of ECG Signals Based on MIT-BIH Dataset Using Bi-LSTM Model for Assisting Cardiologists Diagnosis. *Traitement du Signal*. 41(6): 3245-3253.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 3149–3157.
- Luz, E. J. d. S., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D. (2016). ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer Methods and Programs in Biomedicine*, 127, 144–164. <https://doi.org/10.1016/j.cmpb.2015.12.008>.
- Malghan, P. G., & Hota, M. K. (2020). A review on ECG filtering techniques for rhythm analysis. *Research on Biomedical Engineering*, 36(2), 171-186.
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2), 153–157.
- Moody, G. B., Mark, R. G. (2001). The impact of the MIT-BIH arrhythmia

- database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3), 45–50. <https://doi.org/10.1109/51.932724>.
- Pambudi, R. E., Purnomo, H. and Aglasia, A. (2025). Analisis Klasifikasi Sentimen Pengguna MyPertamina Menggunakan Metode Evaluasi Precision, Recall, dan F1-Score. *Journal of Informatics and Electrical Engineering*. 7(2): 17–22.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830. <https://jmlr.org/papers/v12/pedregosa11a.html>.
- Putra, T. W., Triayudi, A. and Andrianingsih (2022). Analisis Sentimen Pembelajaran Daring menggunakan Metode Naïve Bayes, KNN, dan Decision Tree. *Jurnal Teknologi Informasi dan Komunikasi (JTik)*. 6(1): 20–26.
- Rahul, J., Sora, M., Sharma, L. D., and Bohat, V. K. (2021). An improved cardiac arrhythmia classification using an RR interval-based approach. *Biocybernetics and Biomedical Engineering*, 41(3). <https://doi.org/10.1016/j.bbe.2021.04.004>.
- Sahoo, S., Dash, M., Behera, S., & Sabut, S. (2020). Machine learning approach to detect cardiac arrhythmias in ECG signals: A survey. *Irbm*. 41(4), 185-194.
- van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.
- Zhou, C., Li, X., Feng, F., Zhang, J., Lyu, H., Wu, W., Tang, X., Luo, B., Li, D., Xiang, W. and Yao, D. (2023) Inter-patient ECG heartbeat classification for arrhythmia classification: a new approach of multi-layer perceptron with weight capsule and sequence-to-sequence combination. *Front. Physiol.* 14:1247587. <https://doi.org/10.3389/fphys.2023.1247587>.