

**K-MEANS CLUSTERING UNTUK SEGMENTASI
INTENSITAS AKTIVITAS LARI BERDASARKAN DATA
GPS DAN HEART RATE DI WILAYAH PERKOTAAN
DAN KABUPATEN: STUDI MENGGUNAKAN DATA
SIMULASI**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh

FAISAL BIN MUHAMMAD IMRAN

20.83.0555

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

**K-MEANS CLUSTERING UNTUK SEGMENTASI
INTENSITAS AKTIVITAS LARI BERDASARKAN DATA
GPS DAN HEART RATE DI WILAYAH PERKOTAAN DAN
KABUPATEN: STUDI MENGGUNAKAN DATA SIMULASI
SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Teknik Komputer



disusun oleh

FAISAL BIN MUHAMMAD IMRAN

20.83.0555

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**K-MEANS CLUSTERING UNTUK SEGMENTASI INTENSITAS
AKTIVITAS LARI BERDASARKAN DATA GPS DAN HEART RATE
DI WILAYAH PERKOTAAN DAN KABUPATEN: STUDI
MENGUNAKAN DATA SIMULASI**

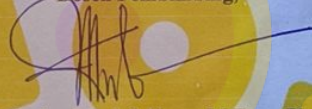
yang disusun dan diajukan oleh

Faisal Bin Muhammad Imran

20.83.0555

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 02 Maret 2026

Dosen Pembimbing,



Melwin Syafrizal, S.Kom., M.Eng., Ph.D
NIK. 190302105

HALAMAN PENGESAHAN

SKRIPSI

K-MEANS CLUSTERING UNTUK SEGMENTASI INTENSITAS
AKTIVITAS LARI BERDASARKAN DATA GPS DAN HEART RATE DI
WILAYAH PERKOTAAN DAN KABUPATEN: STUDI MENGGUNAKAN
DATA SIMULASI

yang disusun dan diajukan oleh

Faisal Bin Muhammad Imran
20.83.0555

Telah dipertahankan di depan Dewan Penguji
pada tanggal 02 Maret 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

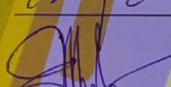
Muhammad Kopravi, S.Kom., M.Eng.
NIK. 190302454



Andika Agus Slemeto, S.Kom., M.Kom
NIK. 190302109



Melwin Syafrizal, S.Kom., M.Eng., Ph.D.
NIK. 190302105



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 02 Maret 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Faisal Bin Muhammad Imran
NIM : 20.83.0555

Menyatakan bahwa Skripsi dengan judul berikut:

**K-MEANS CLUSTERING UNTUK SEGMENTASI INTENSITAS
AKTIVITAS LARI BERDASARKAN DATA GPS DAN HEART RATE DI
WILAYAH PERKOTAAN DAN KABUPATEN: STUDI MENGGUNAKAN
DATA SIMULASI**

Dosen Pembimbing : Melwin Syafrizal, S.Kom., M.Eng., Ph.D

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 02 Maret 2026

Yang Menyatakan,



Faisal Bin Muhammad Imran

HALAMAN PERSEMBAHAN

Karya ini saya persembahkan kepada Allah SWT yang telah memberikan rahmat, hidayah, dan kemudahan dalam setiap langkah perjalanan studi saya. Tanpa henti-hentinya saya panjatkan doanya kepada sang penguasa atas segala hal sehingga bisa menyelesaikan karya ini.

Beribu-ribu rasa terima kasih saya ucapkan kepada kedua orang tua tercinta, ayah dan ibu serta adik tercinta yang selalu bantu memberikan doa, dukungan, dan kasih sayang baik secara lahir maupun batin tiada henti. Semoga karya ini menjadi salah satu bentuk bakti dan kebanggaan bagi kalian.

Tak lupa juga, kepada calon buah hati saya yang turut ikut serta dalam proses penyelesaian karya ini sehingga saya ingin mengucapkan terima kasih sebanyak-banyaknya karena tiada henti menyemangati sampai titik dimana saya menyelesaikan karya ini.

Terakhir, kepada bapak/ibu dosen dan seluruh staf, sahabat dan teman karib yang tidak bisa saya sebutkan satu per satu ilmu, bimbingan, dan dukungan selama masa perkuliahan, semoga seluruh dukungan dibales dengan pahala yang berlipat ganda.

Dengan rasa penuh berterima kasih dan rendah hati, karya ini saya persembahkan.

KATA PENGANTAR

Puji Syukur penulis panjatkan kepada tuhan yang maha esa yang telah melimpahkan rahmat, hidayah, dan karunia-nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “K-Means Clustering untuk Segmentasi Intensitas Aktivitas Lari Berdasarkan Data GPS dan Heart Rate di Wilayah Perkotaan dan Kabupaten: Studi Menggunakan Data Simulasi”. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana komputer pada program studi teknik komputer universitas amikom Yogyakarta.

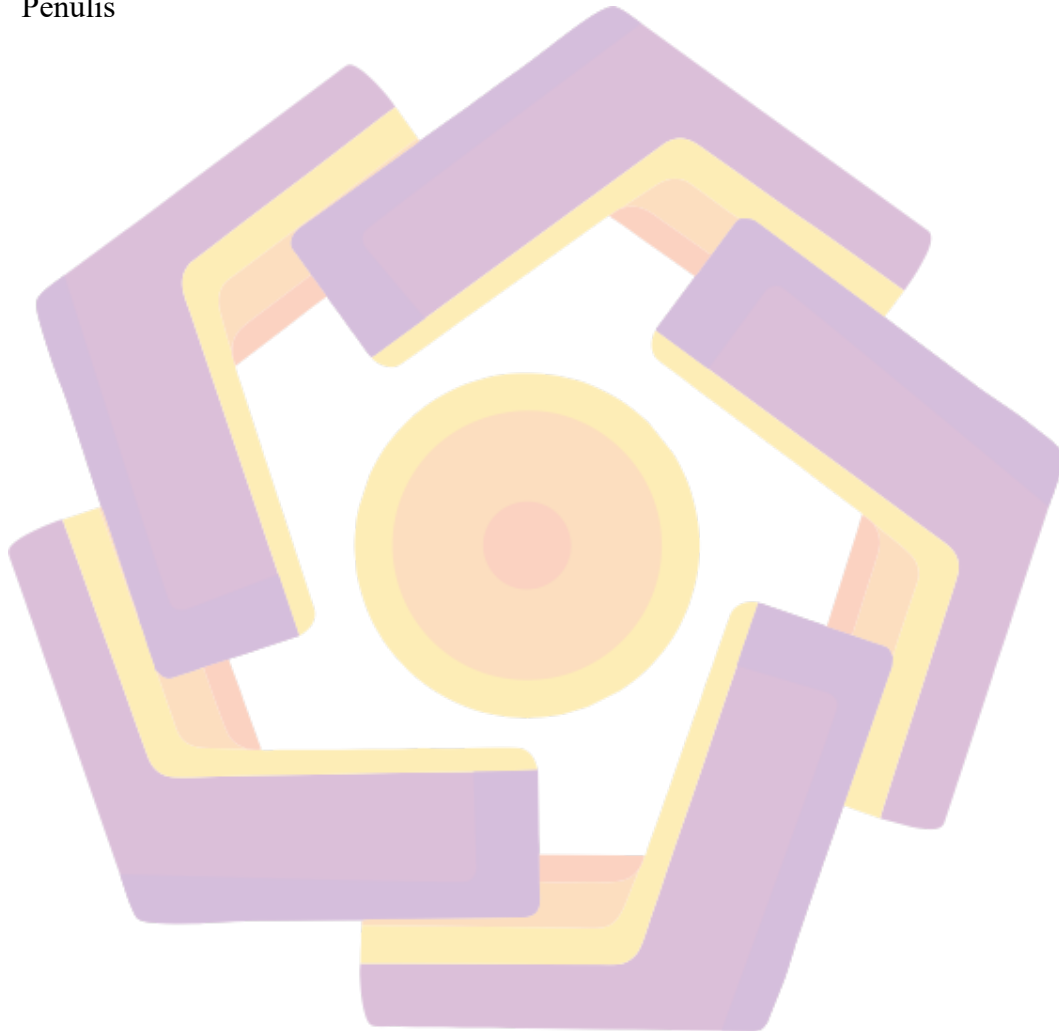
Penulis menyadari bahwa penyelesaian skripsi ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT yang telah memberikan kesehatan, kekuatan, dan kemudahan dalam menyelesaikan skripsi ini.
2. Ayah, Ibu tercinta serta adik tersayang yang selalu memberikan doa, dukungan moral, dan semangat yang tiada henti selama penulis menempuh pendidikan.
3. Reina Sabilatunnajah, S.Hum yang tiada henti memberikan dukungan sebesar-besarnya berupa doa dan motivasi selama penulis menyusun skripsi ini.
4. Bapak Melwin Syafrizal, S.Kom., M.Eng., Ph.D yang telah meluangkan waktu, tenaga, dan pikiran untuk membimbing penulis dengan penuh kesabaran dan memberikan arahan yang sangat berharga dalam penyusunan skripsi ini.
5. Rekan TK02 dan teman-teman seperjuangan yang telah memberikan motivasi, bantuan, dan semangat dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca.

Yogyakarta, 11 Februari 2026

Penulis



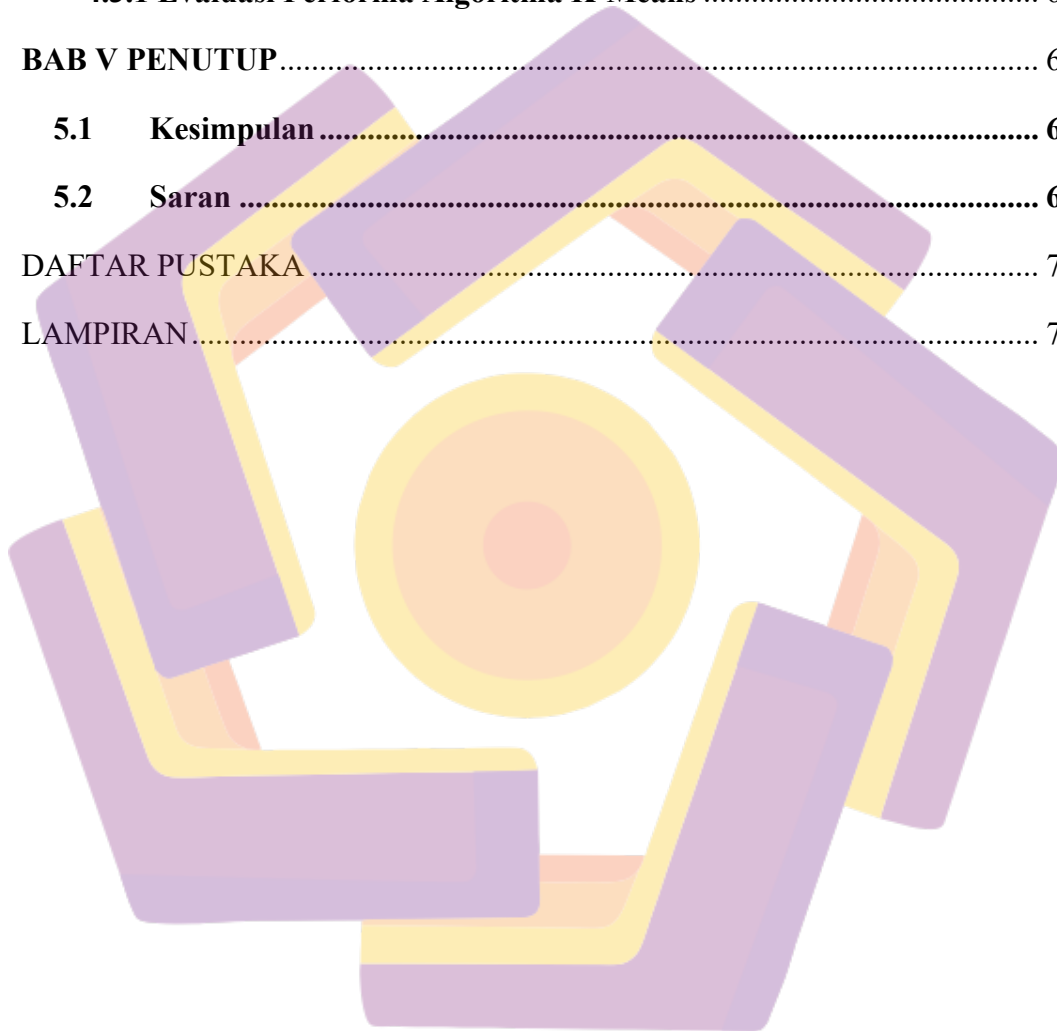
DAFTAR ISI

HALAMAN PERSETUJUAN.....	III
HALAMAN PENGESAHAN.....	IV
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	V
HALAMAN PERSEMBAHAN	VI
KATA PENGANTAR	VII
DAFTAR ISI.....	VIII
DAFTAR TABEL.....	XI
DAFTAR GAMBAR	XII
DAFTAR LAMPIRAN.....	XIII
DAFTAR LAMBANG DAN SINGKATAN	XIV
DAFTAR ISTILAH	XVI
INTISARI.....	XXIV
ABSTRACT.....	XXV
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5

BAB II TINJAUAN PUSTAKA.....	7
2.1 Studi Literatur.....	7
2.1.1 Penelitian Terkait IoT dalam Monitoring Aktivitas Fisik	7
2.1.2 Penelitian Terkait Pemrosesan Data GPS	7
2.1.3 Penelitian Terkait Pemrosesan Data Detak Jantung.....	8
2.1.4 Penelitian Terkait Segmentasi Aktivitas Lari	8
2.1.5 Penelitian Terkait K-Means dalam Aktivitas Fisik dan Data Fisiologi	8
2.1.6 Gap Analisis dan Relevansi Penelitian.....	9
2.2 Dasar Teori	10
2.2.1 Aktivitas Fisik Pelari.....	10
2.2.2 Internet of Things (IoT).....	11
2.2.3 Global Positioning System (GPS)	11
2.2.4 Sensor Detak Jantung.....	12
2.2.5 Pengolahan Data GPS dan Detak Jantung Sebelum Clustering ...	13
2.2.6 Machine Learning	14
2.2.8 Google Colaboratory (Google Colab).....	15
2.2.7 K-Means Clustering.....	16
2.2.9 Bahasa Pemrograman Python	17
2.2.10 Variabel Penelitian.....	18
BAB III METODE PENELITIAN	22
3.1 Objek Penelitian	22

3.2	Alur Penelitian.....	22
3.3	Alat dan Bahan.....	25
3.3.1	Alat Penelitian	26
3.3.2	Bahan Penelitian	27
3.4	Data Penelitian	27
3.5	Metode Analisis Data.....	28
3.5.1	Eksplorasi Data Awal	28
3.6	Preprocessing Data.....	28
3.7	Metode Clustering.....	30
3.7.1	Alasan menggunakan K-Means.....	30
3.7.2	Validasi Cluster.....	30
3.8	Analisis Perbandingan Wilayah	30
BAB IV HASIL DAN PEMBAHASAN		32
4.1	Hasil Clustering.....	32
4.1.1	Penentuan Jumlah Cluster Optimal.....	32
4.1.2	Hasil Clustering dengan K-Means (k=3)	37
4.2	Visualisasi dan Analisis Hasil Clustering.....	43
4.2.1	Analisis Hubungan Speed dan Detak Jantung	43
4.2.2	Analisis Hubungan Pace dan Detak Jantung	45
4.2.3	Distribusi Heart Rate per Cluster	46
4.2.4	Sebaran Geografis GPS.....	49
4.2.5	Profil Elevasi dan Medan.....	53

4.2.6	Kompleksitas Rute (Route Complexity Index)	55
4.2.7	Analisis Korelasi Fitur per Cluster	58
4.3	Pembahasan	61
4.3.1	Evaluasi Performa Algoritma K-Means	61
BAB V	PENUTUP	66
5.1	Kesimpulan	66
5.2	Saran	67
DAFTAR PUSTAKA	72
LAMPIRAN	76



DAFTAR TABEL

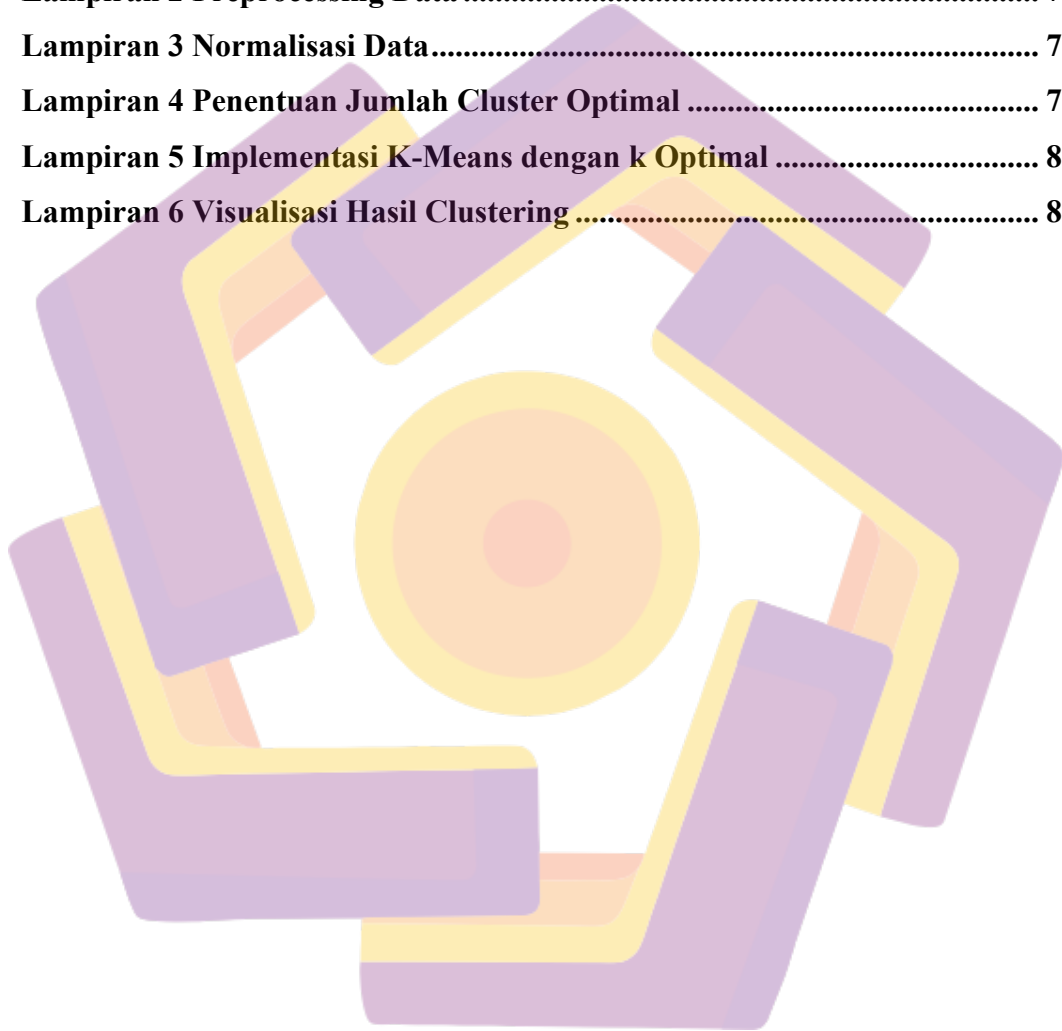
Tabel 2.2 Variabel Penelitian	24
Tabel 3.1 Alat Penelitian	30
Tabel 3.2 Data Penelitian	32
Tabel 4.1 Nilai Inertia untuk Setiap k	33
Tabel 4.2 Silhouette Score untuk Setiap k	35
Tabel 4.3 Distribusi Data Per Cluster	38
Tabel 4.4 Statistik Deskriptif Karakteristik Setiap Cluster	39
Tabel 4.5 Centroid Setiap Cluster (Nilai Ternormalisasi)	39
Tabel 4.6 Metrik Kualitas Clustering	42
Tabel 4.7 Statistik Heart Rate per Cluster	48
Tabel 4.8 Distribusi Cluster per Wilayah (Cross-Tabulation)	51
Tabel 4.9 Hasil Uji ANOVA Satu Arah Antar Cluster	51
Tabel 4.10 Uji Mann-Whitney Perbandingan Fitur Perkotaan vs Kabupaten	52
Tabel 4.11 Analisis Elevasi per Cluster	54
Tabel 4.12 Statistik RCI per Cluster	56
Tabel 4.13 Korelasi Signifikan per Cluster	60
Tabel 4.14 Summary Validasi Internal	61

DAFTAR GAMBAR

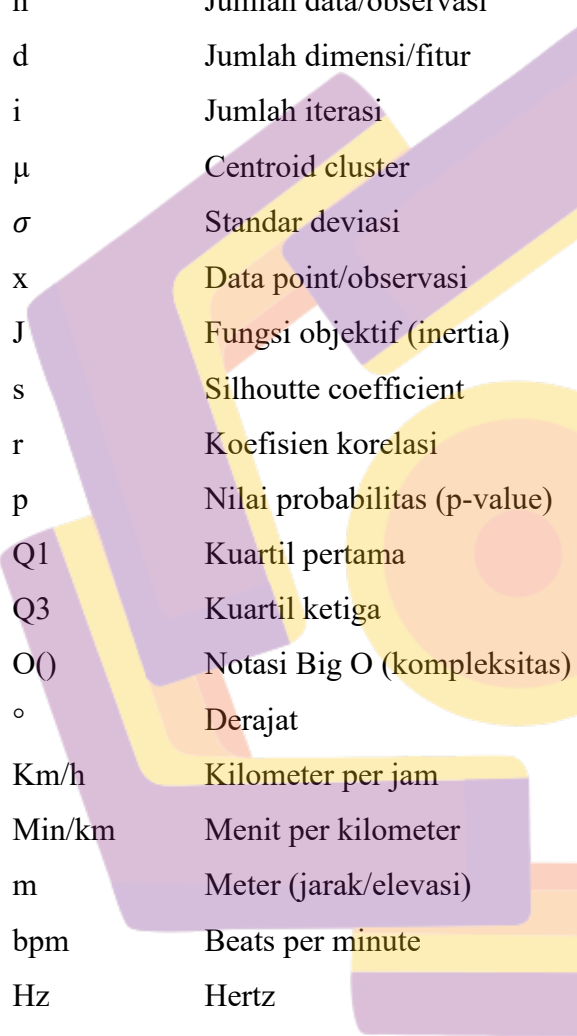
Gambar 2.1 Aktivitas lari.....	15
Gambar 2.2 Sistem IoT.....	16
Gambar 2.3 Global Positioning System	17
Gambar 2.4 Sensor Detak Jantung pada jam tangan pintar	18
Gambar 2.5 Tipe Machine learning	20
Gambar 2.6 Flowchart K-Means Clustering.....	21
Gambar 2.7 Google Colab	22
Gambar 2.8 Logo Python	23
Gambar 3.1 Flowchart Alur Penelitian.....	28
Gambar 4.1 Elbow Method	35
Gambar 4.2 Grafik Silhouette Score	37
Gambar 4.3 Kombinasi Elbow dan Silhouette Score.....	39
Gambar 4.4 Distribusi Proporsi Cluster	40
Gambar 4.5 Profil Karakteristik Multi-Dimensi (Radar Chart)	43
Gambar 4.6 Scatter Plot Speed vs Heart Rate per Cluster	46
Gambar 4.7 Scatter Plot Pace vs Heart Rate per Cluster	48
Gambar 4.8 Boxplot Distribusi Heart Rate per Cluster.....	49
Gambar 4.9 Scatter Plot GPS Track per Cluster.....	52
Gambar 4.10 Profil Elevasi per Cluster	54
Gambar 4.11 Boxplot Route Complexity Index (RCI) per Cluster.....	56
Gambar 4.12 Heatmap Korelasi Fitur per Cluster.....	58

DAFTAR LAMPIRAN

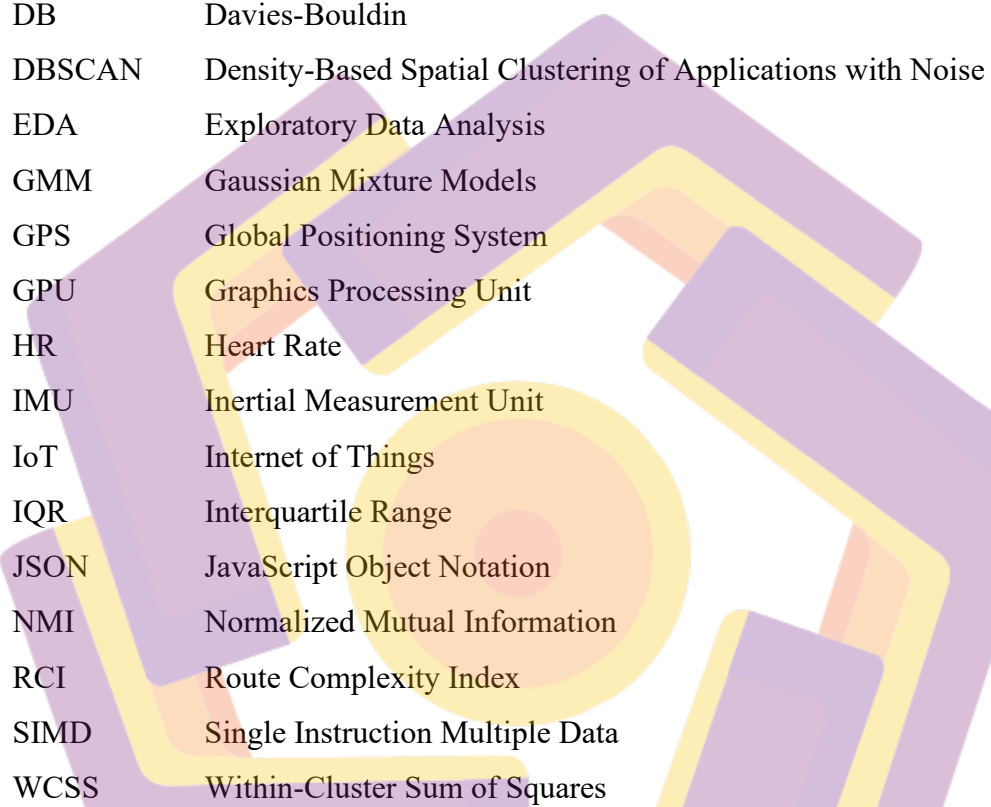
Lampiran 1 Import Libraries dan Load Data.....	75
Lampiran 2 Preprocessing Data	76
Lampiran 3 Normalisasi Data.....	77
Lampiran 4 Penentuan Jumlah Cluster Optimal	79
Lampiran 5 Implementasi K-Means dengan k Optimal	80
Lampiran 6 Visualisasi Hasil Clustering	81



DAFTAR LAMBANG DAN SINGKATAN

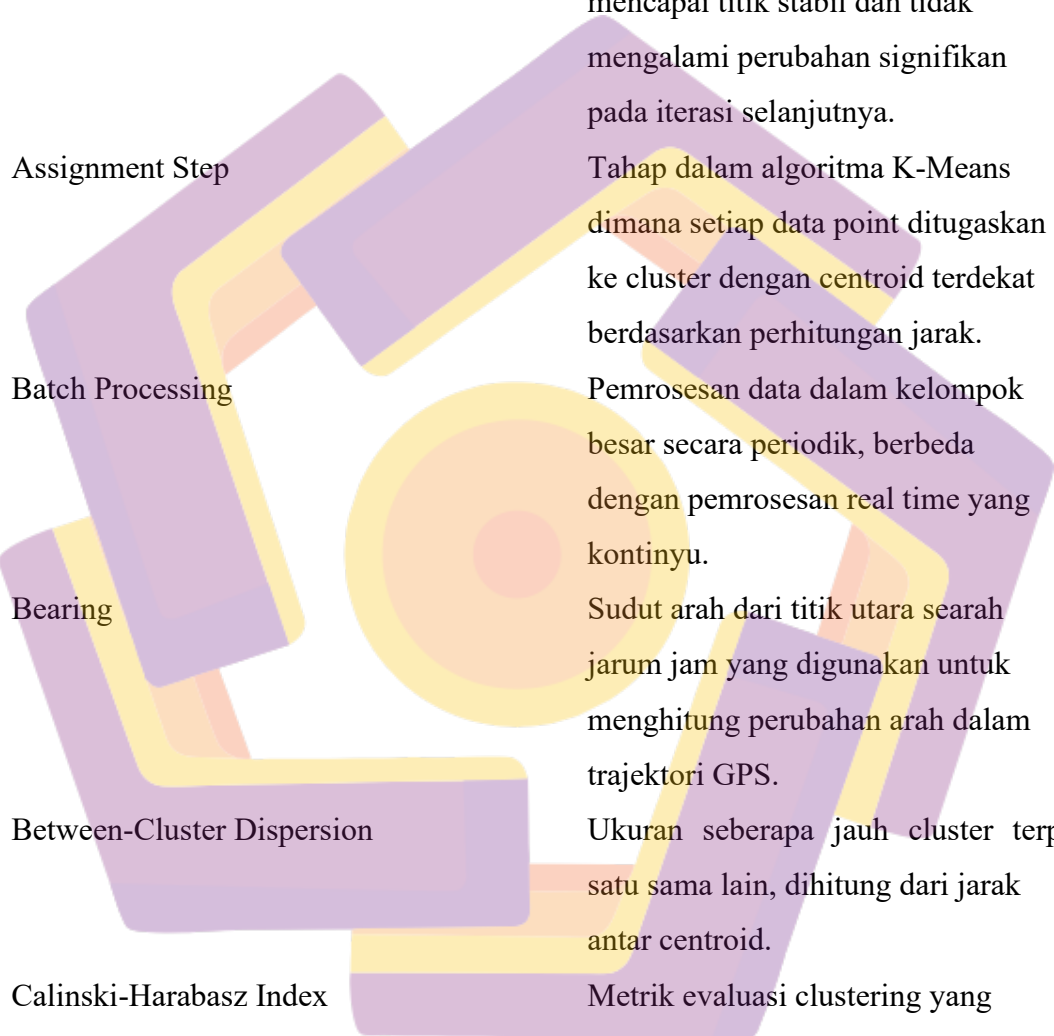


k	Jumlah cluster
n	Jumlah data/observasi
d	Jumlah dimensi/fitur
i	Jumlah iterasi
μ	Centroid cluster
σ	Standar deviasi
x	Data point/observasi
J	Fungsi objektif (inertia)
s	Silhouette coefficient
r	Koefisien korelasi
p	Nilai probabilitas (p-value)
Q1	Kuartil pertama
Q3	Kuartil ketiga
O()	Notasi Big O (kompleksitas)
°	Derajat
Km/h	Kilometer per jam
Min/km	Menit per kilometer
m	Meter (jarak/elevasi)
bpm	Beats per minute
Hz	Hertz
MB	Megabyte
KB	Kilobyte
%	Persen
AI	Artificial Intelligence



API	Application Programming Interface
ARI	Adjusted Rand Index
CI/CD	Continuous Integration/ Continuous Deployment
CSV	Comma-Separated Values
DB	Davies-Bouldin
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
EDA	Exploratory Data Analysis
GMM	Gaussian Mixture Models
GPS	Global Positioning System
GPU	Graphics Processing Unit
HR	Heart Rate
IMU	Inertial Measurement Unit
IoT	Internet of Things
IQR	Interquartile Range
JSON	JavaScript Object Notation
NMI	Normalized Mutual Information
RCI	Route Complexity Index
SIMD	Single Instruction Multiple Data
WCSS	Within-Cluster Sum of Squares

DAFTAR ISTILAH



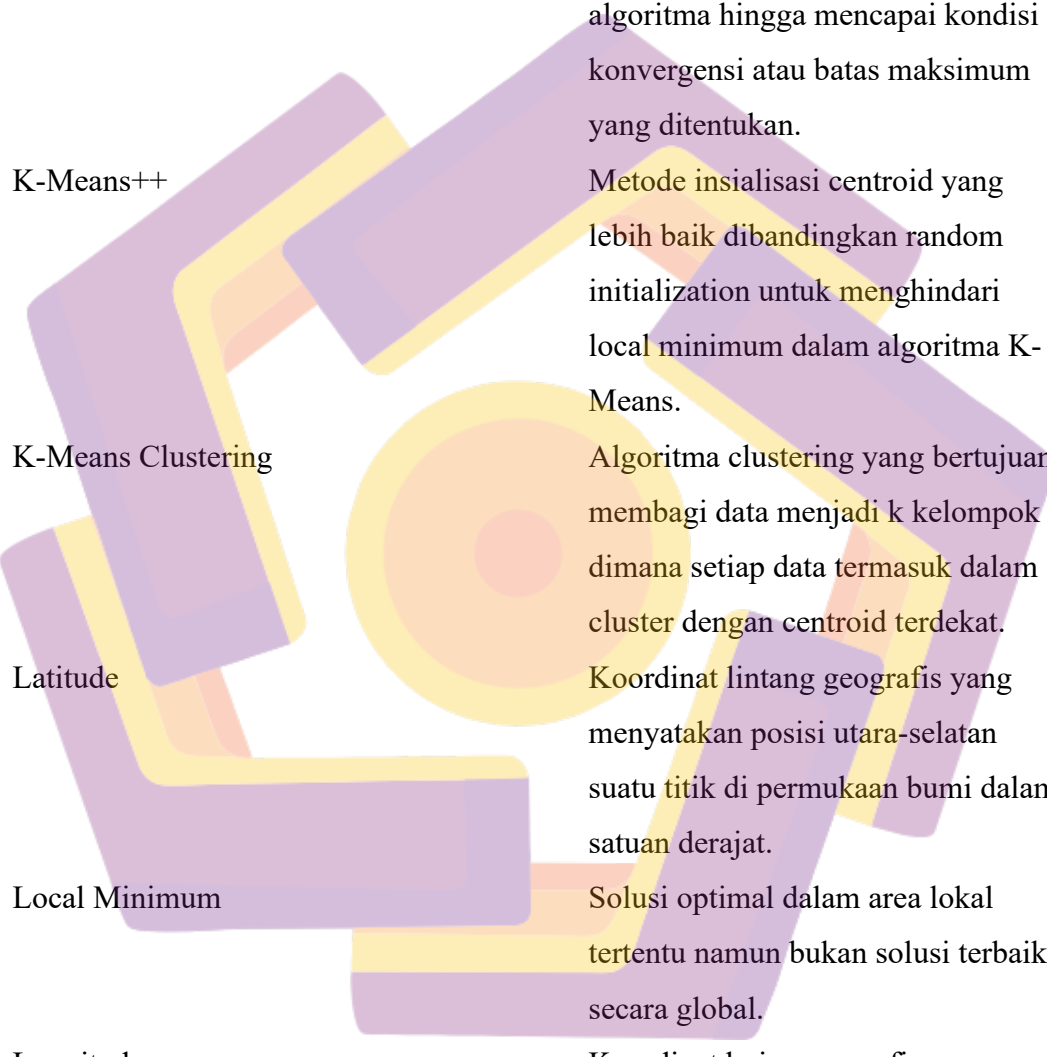
Algorithm Convergence	Kondisi dimana algoritma iterative mencapai titik stabil dan tidak mengalami perubahan signifikan pada iterasi selanjutnya.
Assignment Step	Tahap dalam algoritma K-Means dimana setiap data point ditugaskan ke cluster dengan centroid terdekat berdasarkan perhitungan jarak.
Batch Processing	Pemrosesan data dalam kelompok besar secara periodik, berbeda dengan pemrosesan real time yang kontinyu.
Bearing	Sudut arah dari titik utara searah jarum jam yang digunakan untuk menghitung perubahan arah dalam trajektori GPS.
Between-Cluster Dispersion	Ukuran seberapa jauh cluster terpisah satu sama lain, dihitung dari jarak antar centroid.
Calinski-Harabasz Index	Metrik evaluasi clustering yang mengukur rasio antara between-cluster variance dengan within-cluster variance, dimana nilai lebih

Centroid	tinggi menunjukkan clustering lebih baik.
Clustering	Titik pusat dari suatu cluster yang merepresentasikan rata-rata dari semua data point dalam cluster tersebut.
Convergence Criteria	Proses pengelompokan data menjadi beberapa kelompok berdasarkan kemiripan karakteristik tanpa menggunakan label yang sudah ditentukan sebelumnya.
Data Cleaning	Kriteria atau threshold yang digunakan untuk menentukan kapan algoritma iterative dianggap sudah konvergen dan dapat dihentikan.
Data Dummy	Proses pembersihan data untuk menghilangkan nilai yang tidak valid, tidak konsisten, atau duplikat.
Davies-Bouldin Index	Data simulasi atau sintetis yang dibuat untuk merepresentasikan karakteristik data nyata dengan tujuan demonstrasi atau penelitian.
	Metrik evaluasi clustering yang mengukur rasio within-cluster dispersion terhadap between-cluster separation, dimana nilai lebih rendah menunjukkan clustering lebih baik.

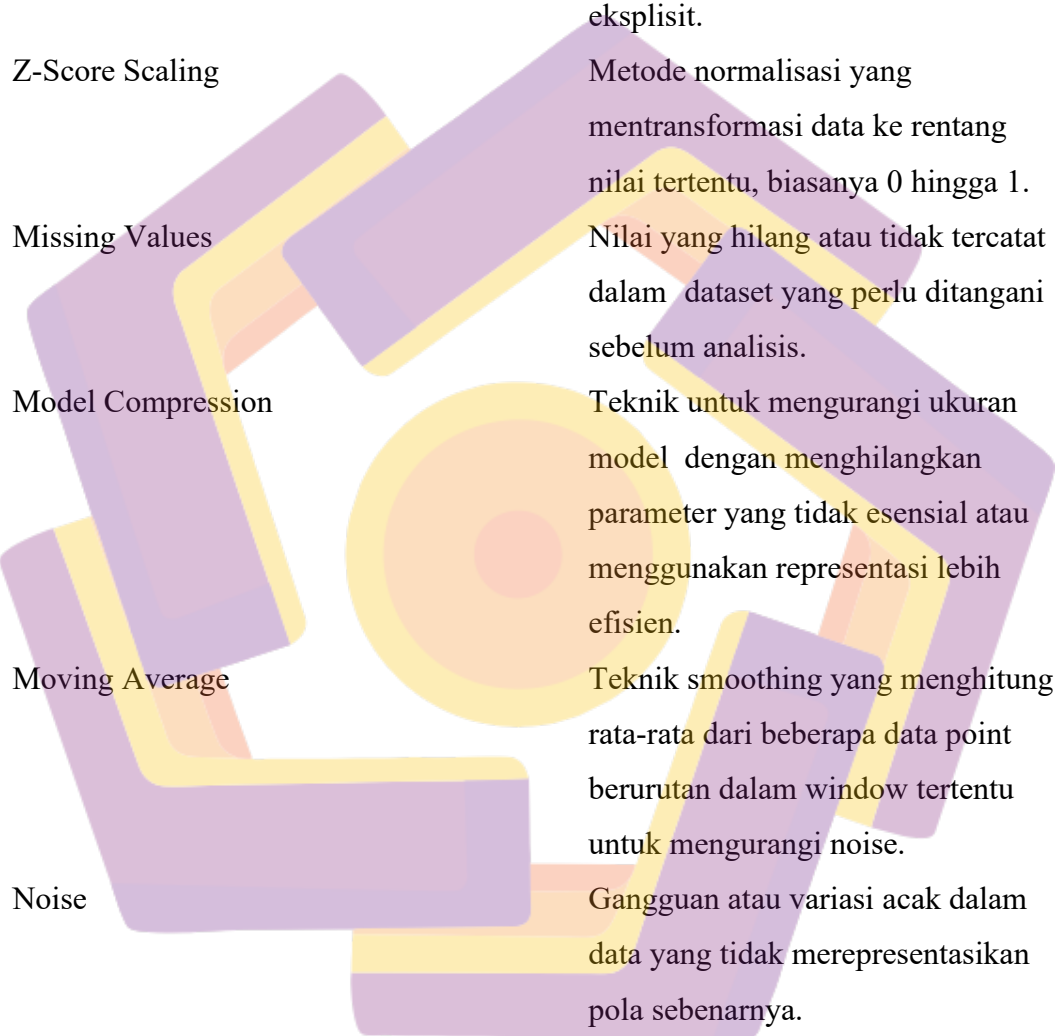


Deployment	Proses implementasi model atau sistem ke lingkungan produksi untuk digunakan secara operasional.
Drift Detection	Teknik untuk mendeteksi perubahan distribusi data input yang dapat mengindikasikan penurunan performa model.
Elbow Method	Metode untuk menentukan jumlah cluster optimal dengan menganalisis penurunan nilai inertia, dimana titik siku menunjukkan jumlah cluster yang tepat.
Elevasi	Ketinggian suatu titik dari permukaan laut yang diukur dalam satuan meter.
Euclidean Distance	Metrik jarak yang mengukur jarak lurus antara dua titik dalam ruang multi-dimensi menggunakan teorema Pythagoras.
Feature Engineering	Proses pembuatan variabel atau fitur baru dari data mentah untuk meningkatkan kualitas analisis atau performa model.
Filtering	Proses penyaringan data untuk menghilangkan noise atau gangguan menggunakan teknik seperti moving average atau Kalman filter.

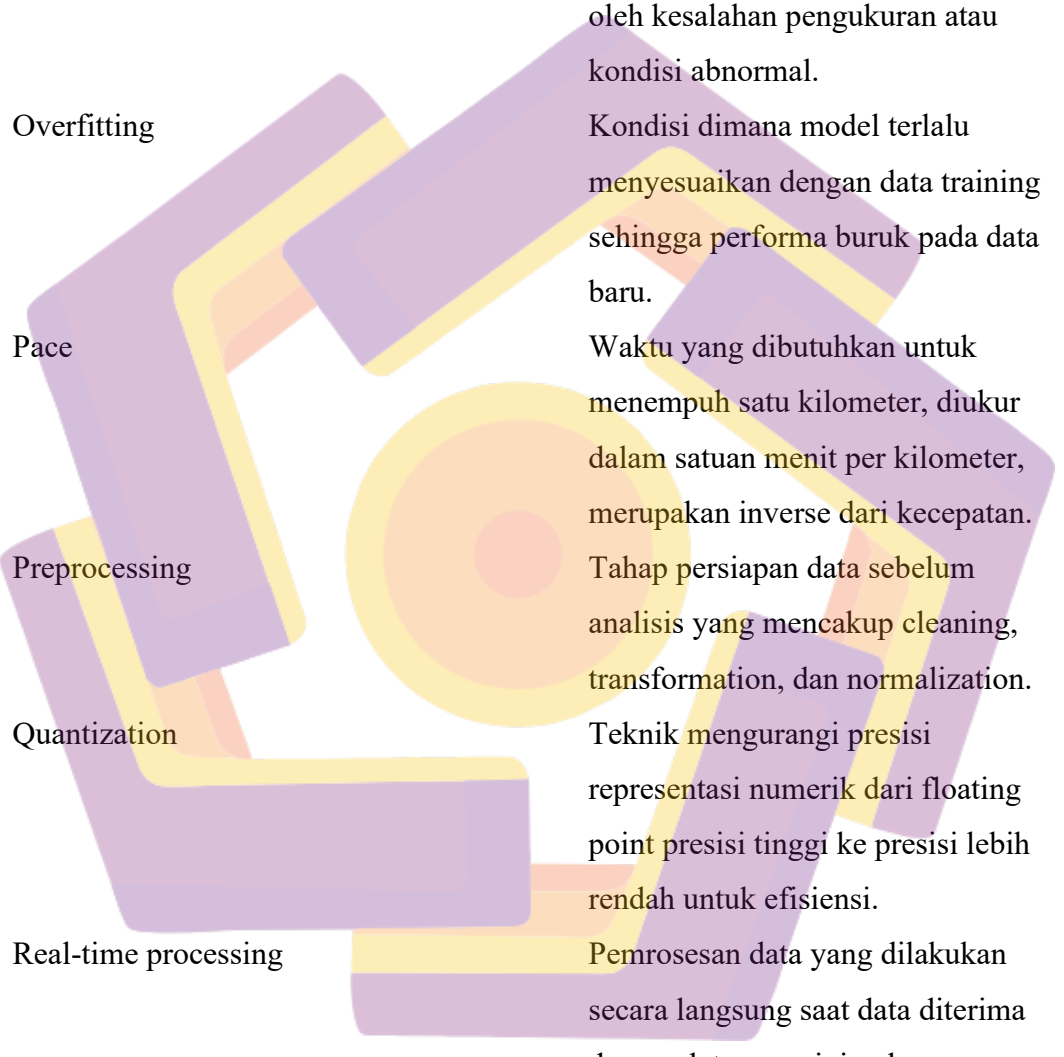
Global Optimum	Solusi terbaik secara keseluruhan dalam ruang solusi, berbeda dengan local optimum yang hanya terbaik di area tertentu.
GPS Drift	Fenomena pergeseran posisi GPS secara acak yang disebabkan oleh gangguan sinyal satelit atau multipath effect.
Haversine Formula	Formula matematis untuk menghitung jarak geodesik antara dua titik koordinat GPS dengan memperhitungkan kelengkungan permukaan bumi.
Heart Rate	Detak jantung yang diukur dalam satuan denyut per menit (bpm) sebagai indikator respons fisiologis tubuh terhadap aktivitas lari.
Homogenitas	Tingkat kesamaan atau konsistensi karakteristik data dalam satu cluster.
Horizontal Scaling	Penambahan kapasitas sistem dengan menambah jumlah instance atau node yang bekerja secara paralel.
Inertia	Ukuran total within-cluster sum of squares yang mengindikasikan seberapa dekat data dalam satu cluster dengan centroid clusternya.



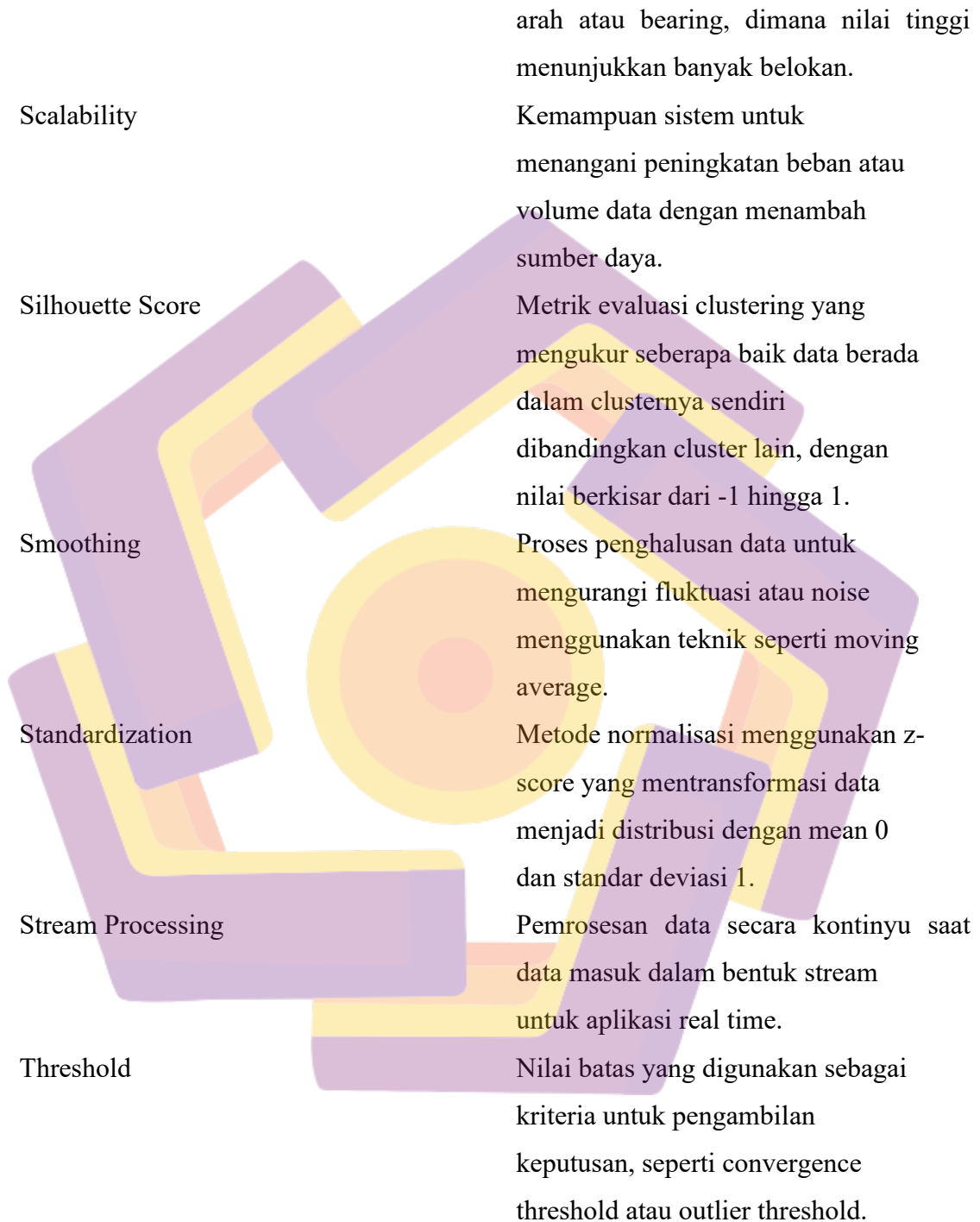
Interpolation	Teknik untuk mengisi nilai yang hilang dengan mengestimasi nilai berdasarkan data di sekitarnya.
Iterasi	Proses pengulangan langkah-langkah algoritma hingga mencapai kondisi konvergensi atau batas maksimum yang ditentukan.
K-Means++	Metode inisialisasi centroid yang lebih baik dibandingkan random initialization untuk menghindari local minimum dalam algoritma K-Means.
K-Means Clustering	Algoritma clustering yang bertujuan membagi data menjadi k kelompok dimana setiap data termasuk dalam cluster dengan centroid terdekat.
Latitude	Koordinat lintang geografis yang menyatakan posisi utara-selatan suatu titik di permukaan bumi dalam satuan derajat.
Local Minimum	Solusi optimal dalam area lokal tertentu namun bukan solusi terbaik secara global.
Longitude	Koordinat bujur geografis yang menyatakan posisi timur-barat suatu titik di permukaan bumi dalam satuan derajat.




Machine Learning	Cabang dari artificial intelligence yang memungkinkan sistem belajar dan meningkatkan performa dari pengalaman tanpa deprogram secara eksplisit.
Z-Score Scaling	Metode normalisasi yang mentransformasi data ke rentang nilai tertentu, biasanya 0 hingga 1.
Missing Values	Nilai yang hilang atau tidak tercatat dalam dataset yang perlu ditangani sebelum analisis.
Model Compression	Teknik untuk mengurangi ukuran model dengan menghilangkan parameter yang tidak esensial atau menggunakan representasi lebih efisien.
Moving Average	Teknik smoothing yang menghitung rata-rata dari beberapa data point berurutan dalam window tertentu untuk mengurangi noise.
Noise	Gangguan atau variasi acak dalam data yang tidak merepresentasikan pola sebenarnya.
Normalisasi	Proses transformasi data agar setiap variabel memiliki skala yang seragam untuk menghindari



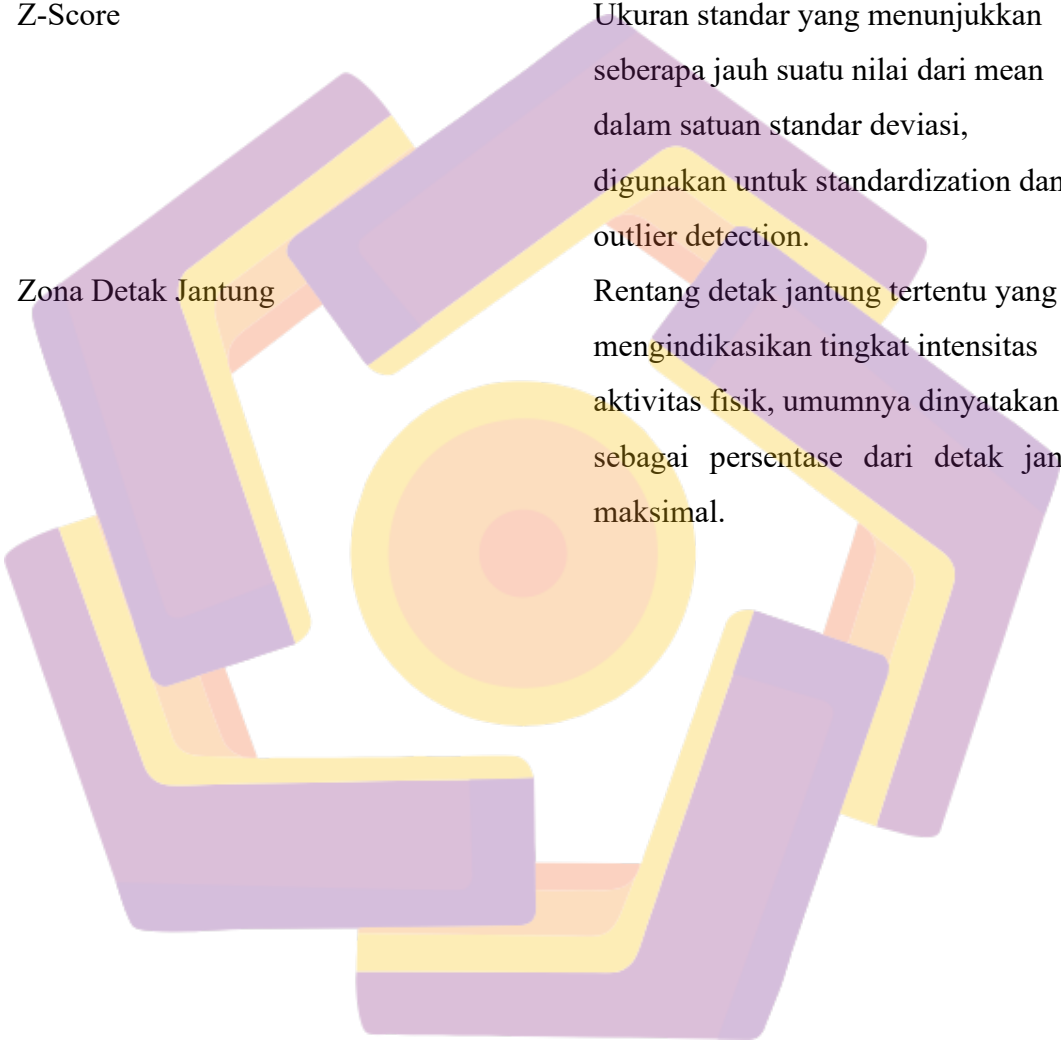
Outlier	<p>dominasi variabel tertentu dalam perhitungan.</p> <p>Nilai yang sangat berbeda dari mayoritas data, dapat disebabkan oleh kesalahan pengukuran atau kondisi abnormal.</p>
Overfitting	<p>Kondisi dimana model terlalu menyesuaikan dengan data training sehingga performa buruk pada data baru.</p>
Pace	<p>Waktu yang dibutuhkan untuk menempuh satu kilometer, diukur dalam satuan menit per kilometer, merupakan inverse dari kecepatan.</p>
Preprocessing	<p>Tahap persiapan data sebelum analisis yang mencakup cleaning, transformation, dan normalization.</p>
Quantization	<p>Teknik mengurangi presisi representasi numerik dari floating point presisi tinggi ke presisi lebih rendah untuk efisiensi.</p>
Real-time processing	<p>Pemrosesan data yang dilakukan secara langsung saat data diterima dengan latency minimal.</p>
Route Complexity Index (RCI)	<p>Indeks yang mengukur kompleksitas rute berdasarkan rata-rata perubahan</p>





Timestamp	Waktu pencatatan data yang digunakan untuk mengurutkan data secara kronologis dan menghitung durasi.
Tolerance	Tingkat toleransi perubahan yang masih dapat diterima sebelum algoritma dianggap konvergen.
Trajektori	Lintasan atau jalur pergerakan yang terekam melalui serangkaian koordinat GPS berurutan.
Underfitting	Kondisi dimana model terlalu sederhana sehingga tidak mampu menangkap pola dalam data dengan baik.
Unsupervised Learning	Jenis machine learning yang mencari pola tersembunyi dalam data tanpa menggunakan label atau kategori yang sudah ditentukan sebelumnya.
Update Step	Tahap dalam algoritma K-Means dimana posisi centroid diperbarui menjadi rata-rata dari semua data point dalam cluster.
Wearable Device	Perangkat yang dapat dikenakan di tubuh seperti smartwatch atau fitness tracker yang dilengkapi sensor untuk monitoring aktivitas.

Within-Cluster Sum of Squares (WCSS)	Total jarak kuadrat antara setiap data point dengan centroid cluster tempat data tersebut berada, digunakan sebagai fungsi objektif K-Means.
Z-Score	Ukuran standar yang menunjukkan seberapa jauh suatu nilai dari mean dalam satuan standar deviasi, digunakan untuk standardization dan outlier detection.
Zona Detak Jantung	Rentang detak jantung tertentu yang mengindikasikan tingkat intensitas aktivitas fisik, umumnya dinyatakan sebagai persentase dari detak jantung maksimal.



INTISARI

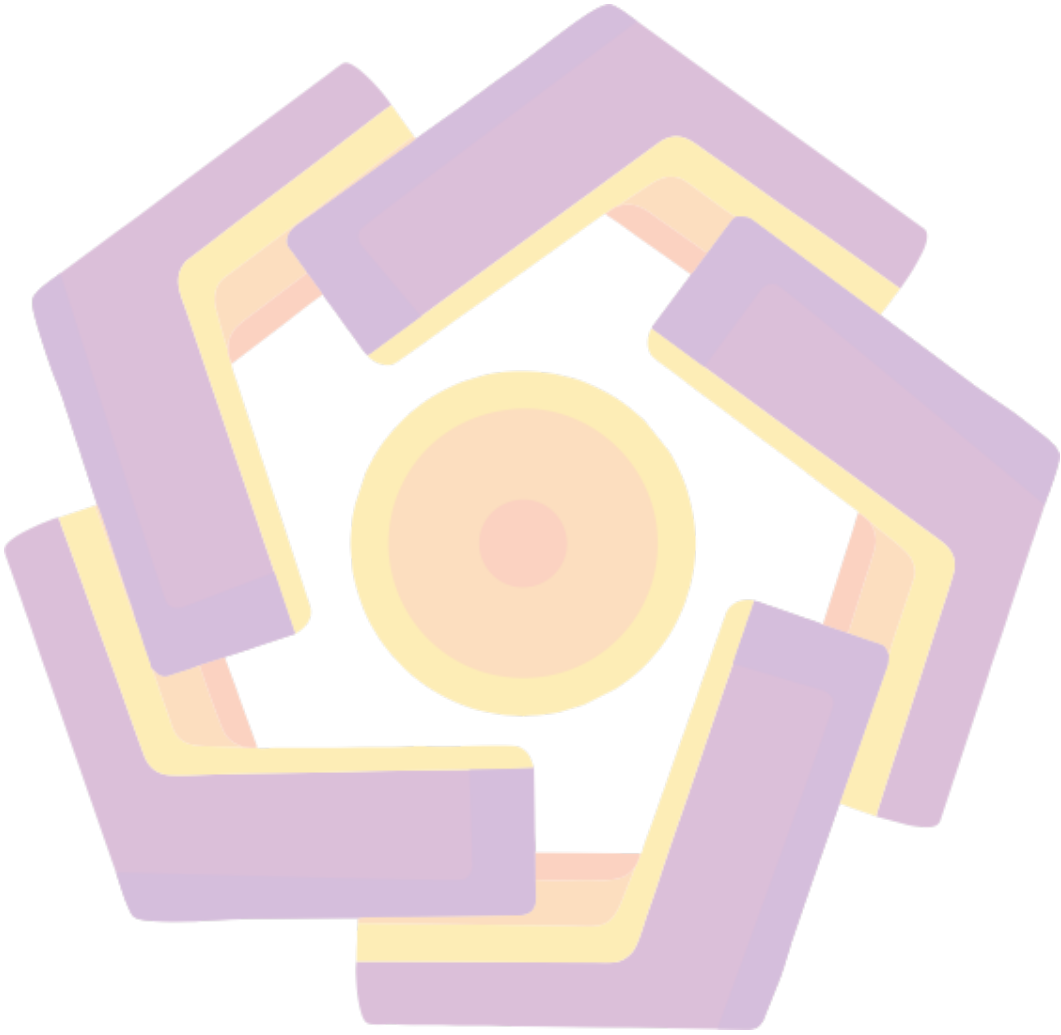
Penggunaan perangkat *wearable* dengan sensor GPS dan detak jantung yang semakin meningkat telah menghasilkan volume data aktivitas yang besar dan memerlukan metode analisis yang efektif untuk segmentasi pola. Penelitian ini bertujuan mengimplementasikan algoritma *K-Means Clustering* untuk melakukan segmentasi pola aktivitas pelari berdasarkan data GPS dan *heart rate* menggunakan data simulasi yang merepresentasikan karakteristik aktivitas lari nyata. Penelitian menggunakan data sekunder berupa data dummy yang merepresentasikan 547 aktivitas lari dengan lima fitur utama yaitu kecepatan, *pace*, detak jantung, elevasi, dan *Route Complexity Index* (RCI).

Metodologi penelitian meliputi *preprocessing* data dengan normalisasi *z-score*, penentuan jumlah *cluster* optimal menggunakan *Elbow Method* dan *Silhouette Score*, serta implementasi *K-Means* dengan inisialisasi *k-means++*. Algoritma mencapai konvergensi pada iterasi ke-47 dengan kompleksitas komputasi $O(n \times k \times i \times d)$. Hasil *clustering* dievaluasi menggunakan tiga metrik yaitu *Silhouette Score* sebesar 0,487, *Davies-Bouldin Index* sebesar 0,823, dan *Calinski-Harabasz Index* sebesar 342,56 yang menunjukkan kualitas *clustering* yang baik dengan pemisahan *cluster* yang jelas dan homogenitas tinggi.

Analisis menghasilkan tiga *cluster* yang memiliki karakteristik berbeda: *Cluster* 0 (36,6%) menunjukkan aktivitas efisiensi kardiovaskular rendah dengan rasio efisiensi detak jantung tinggi, *Cluster* 1 (32,5%) menunjukkan efisiensi kardiovaskular sedang dengan karakteristik bervariasi, dan *Cluster* 2 (30,9%) menunjukkan efisiensi kardiovaskular tinggi dengan efisiensi fisiologis. Normalisasi terbukti meningkatkan performa secara signifikan dengan peningkatan *Silhouette Score* sebesar 56,1 persen.

Penelitian ini memberikan kontribusi pada penerapan *unsupervised learning* untuk analisis IoT dari perangkat *wearable* dan mendemonstrasikan bahwa pendekatan *clustering* dapat memberikan wawasan bermakna untuk segmentasi pola aktivitas.

Kata kunci: *K-Means Clustering, Analisis Data GPS, Monitoring Detak Jantung, Perangkat Wearable, Unsupervised Learning, Analitik IoT.*



ABSTRACT

The increasing use of wearable devices with GPS and heart rate sensors has generated large volume of activity data that require effective analysis methods for pattern segmentation. This study aims to implement the K-Means Clustering algorithm to segment running activity intensity based on GPS and heart rate data using simulated data. The research uses secondary dummy data representing 547 running activity with five main features: speed, pace, heart rate, elevation, and Route Complexity Index (RCI).

The methodology includes data preprocessing with z-score normalization, determination of optimal cluster number using Elbow Method and Silhouette Score, and K-Means implementation with k-means++ initialization. The algorithm achieved convergence at iteration 47 with computational complexity of $O(n \times k \times i \times d)$. Clustering results were evaluated using three metrics: Silhouette Score of 0.487, Davies-Bouldin Index of 0.823, and Calinski-Harabasz Index of 342.56, indicating good clustering quality with clear cluster separation and high homogeneity.

The analysis produced three clusters with distinct characteristics: Cluster 0 (36.6%) representing low-intensity activities with high heart rate efficiency ratio, Cluster 1 (32.5%) representing moderate-intensity with varied characteristics, and Cluster 2 (30.9%) representing high-intensity with physiological efficiency. Normalization proved to significantly improve performance with a 56.1% increase in Silhouette Score. This research contributes to the application of unsupervised learning for IoT data analysis from wearable devices and demonstrates that clustering approaches can provide meaningful insights for activity pattern segmentation.

Keyword: K-Means Clustering, GPS Data Analysis, Heart Rate Monitoring, Wearable Devices, Unsupervised Learning, IoT Analytics