



# Pengelompokan Pola Gerakan Berdasarkan Data Akselerometer Android Menggunakan Metode K-Means untuk Penelitian Pengenalan Aktivitas

## *Grouping Movement Patterns Based on Android Accelerometer Data Using K-Means Method for Activity Recognition Research*

Muhamad Widiyanto<sup>1</sup>, Hasbi Firmansyah<sup>2</sup>

<sup>1</sup> Informatika, Universitas Pancasakti, Tegal, Indonesia, muhammadwidi72@gmail.com

<sup>2</sup> Informatika, Universitas Pancasakti, Tegal, Indonesia, hasbifirmansyah@upstegal.ac.id

### INFORMASI ARTIKEL

#### Sejarah Artikel:

Diterima Redaksi: 02/07/2025

Revisi Akhir: 03/07/2025

Diterbitkan Online: 11/12/2025

### KATA KUNCI

Data Akselerometer, K-means, Pola Gerak

### KORESPONDENSI

Telepon: 0895808953200

### ABSTRAK

Tujuan dari penelitian ini adalah untuk mengidentifikasi pola pergerakan manusia dari data accelerometer menggunakan algoritma K-Means. Data yang digunakan dikumpulkan dari akselerometer ponsel di saku dada, yang mencatat aktivitas berjalan kaki pada rute tertentu. Proses analisis diawali dengan tahap preprocessing yang meliputi pembersihan dan transformasi data, dan diakhiri dengan clustering menggunakan algoritma K-Means. Hasil clustering menunjukkan data terbagi menjadi 5 cluster dengan jumlah data sebanyak 5068 buah. Setiap cluster memiliki jarak rata-rata data ke pusat massanya. Rerata cluster 0 sebesar -97,468, cluster 1 sebesar -112,553, cluster 2 -101,628, cluster 3 -97,306, dan cluster 4 sebesar -96,680. Evaluasi kinerja algoritma menunjukkan bahwa algoritma K-Means efektif dalam mengelompokkan data berdasarkan pola pergerakan dengan rata-rata total jarak -101,042. Kesimpulan dari penelitian ini adalah algoritma K-Means merupakan pendekatan yang andal untuk mengidentifikasi pola pergerakan berdasarkan data akselerometer dan dapat digunakan untuk berbagai aplikasi seperti otentikasi pengguna dan pemantauan aktivitas.

### ABSTRACT

*The objective of this research is to identify human movement patterns from accelerometer data using the K-Means algorithm. The data used was collected from a mobile phone's accelerometer placed in a chest pocket, which recorded walking activity along a specific route. The analysis process began with a preprocessing stage that included data cleaning and transformation, and concluded with clustering using the K-Means algorithm. The clustering results show that the data, consisting of 5068 data points, was divided into 5 clusters. Each cluster has an average distance from its data points to its centroid: cluster 0 at -97.468, cluster 1 at -112.553, cluster 2 at -101.628, cluster 3 at -97.306, and cluster 4 at -96.680. The algorithm's performance evaluation, with a total average distance of -101.042, indicates that the K-Means algorithm is effective in grouping data based on movement patterns. The conclusion of this study is that the K-Means algorithm is a reliable approach for identifying movement patterns from accelerometer data and can be utilized for various applications, such as user authentication and activity monitoring.*

## 1. PENDAHULUAN

Kemajuan teknologi sensor pada perangkat pintar seperti ponsel pintar telah memungkinkan pengumpulan banyak data untuk berbagai tujuan penelitian, seperti mendeteksi aktivitas manusia. Salah satu jenis data yang umum digunakan dalam penelitian ini adalah data

accelerometer. Akselerometer adalah sensor yang dapat mendeteksi perubahan kecepatan dalam berbagai arah, memberikan informasi yang relevan untuk menganalisis pola pergerakan manusia.

Dataset yang digunakan dalam penelitian ini berasal dari ponsel Android yang dibawa di saku dada. Data

akselerometer dikumpulkan dari 22 peserta yang berjalan sepanjang rute yang telah ditentukan. Kumpulan data ini, dengan frekuensi pengambilan sampel "DELAY\_FASTEST" dan konektivitas jaringan dimatikan, dirancang untuk tujuan deteksi aktivitas dan menimbulkan tantangan dalam mengidentifikasi dan mengautentikasi orang berdasarkan pola pergerakan mereka.

K-means, sebuah algoritma pengelompokan dalam penambangan data, digunakan dalam penelitian ini untuk mengelompokkan pola pergerakan yang diperoleh dari data akselerometer. Tidak seperti metode klasifikasi yang memerlukan data berlabel, K-means mengelompokkan data ke dalam beberapa klaster berdasarkan kesamaan fitur, sehingga menjadikannya alat yang berguna untuk mempelajari pola pergerakan tanpa perlu label awal.

Tujuan dari penelitian ini adalah menggunakan data akselerometer untuk mengidentifikasi pola mobilitas peserta yang berbeda dan mengklasifikasikan aktivitas mereka. Studi ini bertujuan untuk memberikan wawasan baru tentang tantangan dan peluang penggunaan pola gerak untuk pengenalan aktivitas dan identifikasi orang berdasarkan data sensor dengan menggunakan metode K-Means. Hasil penelitian ini juga diharapkan dapat menjadi dasar pengembangan lebih lanjut sistem otentikasi berdasarkan pola pergerakan.

## 2. TINJAUAN PUSTAKA

### 2.1. Teori-teori Dasar dan Umum

#### 2.1.1. Data Mining

Penambangan data melibatkan penggalian pola dan informasi yang bermakna dari kumpulan data yang besar dan kompleks. Proses ini melibatkan berbagai teknik seperti clustering, klasifikasi, dan asosiasi untuk memperoleh wawasan yang relevan. Penambangan data adalah proses mengubah data tidak terstruktur menjadi informasi yang dapat ditafsirkan dan berguna dalam mengambil keputusan.

#### 2.1.2. Clustering

*Clustering* merupakan suatu teknik data mining yang bertujuan untuk mengelompokkan data ke dalam kelompok-kelompok (cluster) berdasarkan kesamaan atau kedekatan fitur antar data. Berbeda dengan klasifikasi yang memerlukan data berlabel, clustering merupakan metode pembelajaran tanpa pengawasan yang tidak memerlukan label selama proses analisis data. Clustering mengelompokkan data yang memiliki karakteristik serupa ke dalam cluster yang sama dan menempatkan data dengan karakteristik berbeda ke dalam cluster yang berbeda (Han & Kamber, 2006).

#### 2.1.3. K-means

K-Means adalah algoritma clustering yang paling umum digunakan untuk mengelompokkan data ke dalam jumlah cluster

(k) tertentu. Algoritma ini meminimalkan jarak antar data dalam suatu cluster dan memaksimalkan perbedaan antar cluster. K-means menggunakan pendekatan iteratif untuk menentukan lokasi pusat cluster (centroid) dan mengelompokkan data berdasarkan jarak terdekat dari centroid (MacQueen, 1967).

Langkah-langkah utama algoritma K-Means adalah sebagai berikut:

- Inisialisasi Jumlah Cluster: Menentukan jumlah cluster (k) yang akan digunakan.
- Penentuan Centroid Awal: Memilih titik awal sebagai centroid awal untuk setiap cluster.
- Pengelompokan Data: Setiap data dihitung jaraknya ke centroid dan dikelompokkan ke cluster dengan jarak terdekat.
- Pembaruan Centroid: Posisi centroid diperbarui dengan menghitung rata-rata posisi data dalam setiap cluster.
- Iterasi: Proses pengelompokan dan pembaruan centroid diulangi hingga tidak ada perubahan signifikan dalam posisi centroid atau pengelompokan data.

Rumus K-Means:

Algoritma K-Means menggunakan jarak Euclidean untuk menghitung kedekatan data terhadap centroid. Rumusnya adalah:

$$d_{Euclidean}(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

Gambar 1. Rumus K-means

Keterangan:

$d_{Euclidean}$ : Jarak Euclidean antara data .

$x_i$ : Nilai fitur ke- $i$  dari data  $x$ .

$y_i$ : Nilai fitur ke- $i$  dari data  $y$ .

$i$ : Indeks fitur, dengan  $i = 1, 2, \dots, n$ .

$n$ : Jumlah total fitur pada data. di bawah

Tahapan dalam K-Means:

- Inisialisasi: Menentukan nilai awal untuk centroid secara acak atau berdasarkan metode tertentu.
- Penghitungan Jarak: Menggunakan rumus Euclidean untuk menghitung jarak data ke centroid.
- Pengelompokan: Mengelompokkan data berdasarkan centroid terdekat.
- Pembaruan Centroid: Menghitung ulang centroid berdasarkan rata-rata dari data yang tergabung dalam cluster.
- Evaluasi: Mengevaluasi perubahan dalam pengelompokan. Jika tidak ada perubahan signifikan, proses dihentikan.
- Hasil Akhir: Data dikelompokkan ke dalam cluster, dan setiap cluster mewakili pola tertentu.

2.1.4. *Preprocessing-Data*

*Preprocessing* data adalah tahap penting dalam data mining yang bertujuan untuk mempersiapkan data mentah sehingga siap digunakan dalam proses analisis. Proses ini mencakup beberapa langkah utama, yaitu:

- a. Data Cleaning: Menghilangkan data yang hilang (missing values), memperbaiki data yang tidak konsisten, dan memastikan data bersih dari duplikasi.
- b. Data Transformasi: Mengonversi data mentah ke dalam bentuk yang lebih sesuai untuk analisis, seperti normalisasi atau standarisasi data untuk mengurangi skala yang berbeda antar fitur.
- c. Data Reduksi: Menghapus atribut atau fitur yang tidak relevan sehingga mengurangi ukuran dataset dan meningkatkan efisiensi analisis. Proses ini membantu mengurangi kompleksitas data tanpa kehilangan informasi penting.

**Tabel 1.** Data Set Hasil Preprocessing

NO	Waktu	Akselerasi Sumbu X	Akselerasi Sumbu Y	Akselerasi Sumbu Z
1	0	0.69464	3.1735	7.5048
2	0.030639	0.14982	3.4868	9.2755
3	0.069763	-0.29965	1.9477	9.112
NO	Waktu	Akselerasi Sumbu X	Akselerasi Sumbu Y	Akselerasi Sumbu Z
4	0.099823	-1.6889	1.4165	10.12
.....	....	.....	.....	.....
5069	154.2	-0.46309	4.7944	8.6625

(Sumber: Uci Repository (User Identification From Walking Activity - Pierluigi Casale))

**3. METODOLOGI**

Penelitian ini di lakukan dengan memasukan data yang diambil dari website UCI repository dan di kelompokkan cluster menggunakan aplikasi rapidminer.

**4. HASIL DAN PEMBAHASAN**

**4.1. Retrieve Data**

Tahapan pertama adalah mengambil dataset yang akan digunakan untuk proses clustering. Dataset ini diambil dari file yang telah disiapkan sebelumnya dan berisi data mentah dari sensor akselerometer ponsel Android. Proses ini memastikan bahwa data tersedia dan siap untuk diproses lebih lanjut.

**4.2. Clustering dengan K-Means**

Data yang terkumpul diolah menggunakan algoritma K-Means. Pada penelitian ini kami menggunakan algoritma K-Means untuk mengelompokkan pola pergerakan berdasarkan data accelerometer. Algoritma ini bekerja dengan membagi data menjadi sekumpulan cluster berdasarkan kedekatan fitur. Fase ini

menghasilkan terbentuknya cluster-cluster yang mewakili pola pergerakan peserta penelitian.

**4.3. Evaluasi Kinerja Clustering (Performance)**

Setelah cluster terbentuk, langkah berikutnya adalah mengevaluasi kinerja clustering. Evaluasi ini dilakukan untuk mengukur seberapa baik data dikelompokkan ke dalam cluster masing-masing. Berbagai metrik kinerja seperti silhouette score dan intra-cluster distance digunakan untuk menilai kualitas pengelompokan.

**4.4. Iterasi dan evaluasi**

Setelah cluster terbentuk, langkah berikutnya adalah mengevaluasi kinerja clustering. Evaluasi ini dilakukan untuk mengukur seberapa baik data dikelompokkan ke dalam cluster masing-masing. Berbagai metrik kinerja seperti silhouette score dan intra-cluster distance digunakan untuk menilai kualitas pengelompokan.

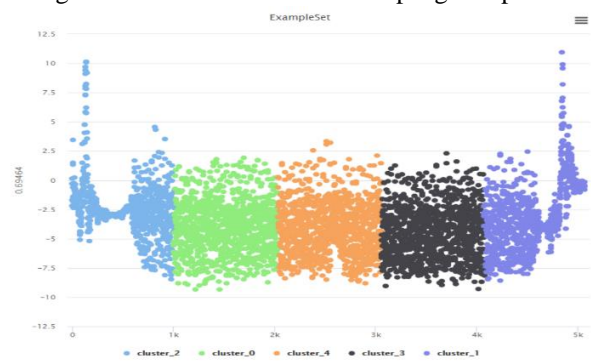
**PerformanceVector**

```
PerformanceVector:
Avg. within centroid distance: -101.042
Avg. within centroid distance_cluster_0: -97.468
Avg. within centroid distance_cluster_1: -112.553
Avg. within centroid distance_cluster_2: -101.628
Avg. within centroid distance_cluster_3: -97.306
Avg. within centroid distance_cluster_4: -96.680
Davies Bouldin: -0.595
```

Gambar 2. Performa Iterasi Dan Evaluasi

**4.5. Visualisasi**

Setelah cluster terbentuk, langkah berikutnya adalah mengevaluasi kinerja clustering. Evaluasi ini dilakukan untuk mengukur seberapa baik data dikelompokkan ke dalam cluster masing-masing. Berbagai metrik kinerja seperti silhouette score dan intra-cluster distance digunakan untuk menilai kualitas pengelompokan.



Gambar 3. Hasil Clustering

**5. KESIMPULAN DAN SARAN**

Penelitian ini menggunakan algoritma K-Means untuk mengelompokkan data berdasarkan pola gerakan akselerometer, dengan hasil yang menunjukkan pengelompokan data yang merata ke dalam lima cluster. Proses iterasi pada K-Means berhasil mencapai konvergensi dengan distribusi data yang cukup seimbang:

- a. Cluster 0: 1.028 items
- b. Cluster 1: 987 items
- c. Cluster 2: 1.009 items

d. Cluster 3: 1.021 items

e. Cluster 4: 1.023 items

Analisis kinerja menunjukkan rata-rata jarak antar data dengan centroidnya, yang mencerminkan kualitas pengelompokan. Cluster 4 memiliki distribusi data paling terpusat, sedangkan Cluster 1 menunjukkan penyebaran data yang lebih luas. Hal ini memberikan wawasan mengenai karakteristik setiap cluster dan bagaimana pola dalam data tersebut teridentifikasi.

Secara keseluruhan, algoritma K-Means terbukti efektif untuk menangkap pola dalam dataset ini, mendukung tujuan penelitian untuk mengidentifikasi dan mengelompokkan pola gerakan berdasarkan data akselerometer. Hasil ini dapat menjadi dasar untuk analisis lanjutan, seperti pengenalan aktivitas atau autentikasi berbasis pola gerakan.

## DAFTAR PUSTAKA

- [1] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. San Francisco, CA: Morgan Kaufmann, 2011.
- [2] J. MacQueen, "Metode untuk Klasifikasi dan Analisis Pengamatan Multivariat," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1. Berkeley, CA: University of California Press, 1967, pp. 281–297.
- [3] P.-N. Tan, M. Steinbach, and V. Kumar, *Pengantar Data Mining*. Boston, MA: Pearson Education, 2005.
- [4] RapidMiner, "Clustering and Performance Evaluation." [Online]. Available: <https://docs.rapidminer.com>. [Diakses: 21 Desember 2024].
- [5] C. C. Aggarwal, *Data Mining: Buku Teks*. New York, NY: Springer, 2015.
- [6] D. T. Larose, *Menemukan Pengetahuan dalam Data: Pengantar Data Mining*. Hoboken, NJ: John Wiley & Sons, 2005.