



# IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING DAN RANDOM FOREST UNTUK PREDIKSI STATUS KESEHATAN KARYAWAN

**Ridwan<sup>1,\*</sup>**

Informatika, Universitas Teknologi Digital

Jln. Cibogo indah III Bodogol, Jawa Barat, Indonesia

Email : [ridwan.gate@gmail.com](mailto:ridwan.gate@gmail.com), [ridwan20123030@digitechuniversity.ac.id](mailto:ridwan20123030@digitechuniversity.ac.id)

Diterima: 23 Juni 2025. Disetujui : 20 Agustus 2025. Dipublikasikan : 05 Desember 2025.

**ABSTRACT** - Employee health is a crucial factor in organizational productivity. This study implements a hybrid data mining approach by combining the K-Means Clustering and Random Forest algorithms to predict employee health status based on daily physical activity data. The dataset consists of 5,000 employee records with 12 attributes, including demographic information, temporal data, and daily step counts. The research methodology utilizes Orange 3 with an integrated workflow for clustering and classification analysis. K-Means clustering successfully identified four distinct employee clusters based on physical activity patterns: Cluster 1 (4,323 steps/day), Cluster 2 (3,478 steps/day), Cluster 3 (9,702 steps/day), and Cluster 4 (4,550 steps/day), with a silhouette score of 0.5, indicating moderate clustering quality with some overlap between clusters. The Random Forest model demonstrated superior performance with an accuracy of 96.4%, AUC of 0.994, F1-score of 0.964, precision of 96.4%, recall of 96.4%, and MCC of 0.934, outperforming Naive Bayes across all evaluation metrics. The integration of both algorithms

provides a comprehensive framework capable of describing existing patterns, predicting future health statuses, and offering specific intervention recommendations for each employee segment. This study contributes practical insights in the form of a cost-effective predictive system for corporate wellness programs, as well as theoretical contributions by validating the effectiveness of ensemble algorithms in the domain of occupational health

**Keywords** : K-Means Clustering, Random Forest, physical Activity, health statuses, data mining.

**ABSTRAK** - Kesehatan karyawan merupakan faktor krusial dalam produktivitas organisasi. Penelitian ini mengimplementasikan pendekatan hybrid data mining yang menggabungkan algoritma K-Means Clustering dan Random Forest untuk memprediksi status kesehatan karyawan berdasarkan data aktivitas fisik harian. Dataset yang digunakan terdiri dari 5.000 record data karyawan dengan 12 atribut meliputi informasi demografis, data temporal, dan jumlah langkah harian. Metodologi penelitian menggunakan

Orange 3 dengan workflow terintegrasi untuk analisis clustering dan klasifikasi. Hasil analisis K-Means clustering berhasil mengidentifikasi 4 cluster karyawan yang distinct berdasarkan pola aktivitas fisik: Cluster 1 (4.323 langkah/hari), Cluster 2 (3.478 langkah/hari), Cluster 3 (9.702 langkah/hari), dan Cluster 4 (4.550 langkah/hari) dengan silhouette score 0,5 yang menunjukkan kualitas clustering yang sedang dengan adanya beberapa tumpang tindih antar cluster. Model Random Forest menunjukkan performa superior dengan akurasi 96,4%, AUC 0,994, F1-score 0,964, precision 96,4%, recall 96,4%, dan MCC 0,934, mengungguli Naive Bayes pada semua metrik evaluasi. Integrasi kedua algoritma menghasilkan framework komprehensif yang mampu mendeskripsikan pola existing, memprediksi status kesehatan dimasa depan, dan memberikan rekomendasi intervensi spesifik per segment karyawan. Penelitian ini memberikan kontribusi praktis berupa sistem prediktif yang cost-effective untuk corporate wellness programs dan kontribusi teoritis dalam validasi efektivitas algoritma ensemble pada domain kesehatan occupational.

**Kata Kunci:** K-Means Clustering, Random Forest, aktivitas fisik, status kesehatan, data mining.

## 1. PENDAHULUAN

Kesehatan karyawan merupakan aspek krusial dalam manajemen sumber daya manusia modern. Organisasi yang proaktif dalam menjaga kesehatan karyawan tidak hanya meningkatkan produktivitas, tetapi juga mengurangi tingkat absensi dan biaya kesehatan [1]. Aktivitas fisik harian, seperti jumlah langkah yang dilakukan, dapat menjadi indikator penting dalam menilai status kesehatan individu. Dengan kemajuan teknologi, data aktivitas fisik kini dapat dikumpulkan secara realtime melalui perangkat wearable, memungkinkan analisis yang lebih mendalam terhadap pola kesehatan karyawan.

PT. Cerebrum Eduknesia Nusantara adalah salah satu perusahaan yang bergerak di bidang

edukasi yang berdomisili di Bandung. Perusahaan setiap tahunnya melakukan evaluasi kesehatan dengan melihat data fisik yaitu langkah. Data yang didapat dari setiap karyawan yang memasukkan data ke aplikasi manakor di setiap harinya. Dan dari data ini peneliti membantu mengidentifikasi apakah karyawan-karyawan diperusahaan sehat atau tidak sehat.

Data mining, khususnya teknik clustering dan classification, telah banyak diterapkan dalam bidang kesehatan. Algoritma K-Means Clustering efektif dalam mengidentifikasi pola tersembunyi dalam data besar [2], sementara Random Forest menunjukkan performa tinggi dalam prediksi status kesehatan berdasarkan data aktivitas fisik [3].

Penelitian ini bertujuan untuk mengimplementasikan kedua algoritma tersebut dalam menganalisis data aktivitas fisik harian karyawan guna memprediksi status kesehatan mereka. Dengan demikian, perusahaan dapat mengambil langkah preventif untuk meningkatkan kesejahteraan karyawan dan produktivitas organisasi.

## 2. TINJAUAN PUSTAKA

### 2.1 Data Mining dalam Kesehatan

Data mining merupakan proses ekstraksi pola dan pengetahuan dari dataset besar menggunakan teknik statistik, machine learning, dan sistem basis data. Dalam domain kesehatan, data mining telah menunjukkan potensi besar untuk mendukung pengambilan keputusan klinis, prediksi penyakit, dan optimasi layanan kesehatan [4]

Penerapan data mining dalam kesehatan kerja khususnya telah mengalami perkembangan signifikan. Implementasi K-Means Clustering dapat mengoptimalkan pengelolaan sumber daya kesehatan dengan mengidentifikasi pola kebutuhan

yang berbeda-beda [1]. Penelitian tersebut mengkonfirmasi bahwa analisis clustering mampu memberikan insights yang berharga untuk perencanaan strategis dalam manajemen kesehatan.

## 2.2 Algoritma K-Means Clustering

K-Means Clustering merupakan salah satu algoritma unsupervised learning yang paling populer untuk membagi data menjadi kluster berdasarkan kemiripan. Algoritma ini bekerja dengan menginisialisasi  $k$  centroid secara acak, kemudian secara iteratif mengassign setiap data point ke centroid terdekat dan mengupdate posisi centroid berdasarkan rata-rata anggota cluster.

Sebagai contoh, penelitian oleh Nisa dan Basir mengaplikasikan K-Means untuk menganalisis distribusi tenaga kesehatan di Provinsi Banten [5]. Selain itu, K-Means juga digunakan untuk mengelompokkan kinerja karyawan berdasarkan pencapaian kerja [6]

## 2.3 Algoritma Random Forest

Random Forest adalah algoritma ensemble learning yang menggabungkan beberapa decision tree untuk meningkatkan akurasi prediksi. Dalam bidang kesehatan, Random Forest telah digunakan untuk memprediksi status kesehatan mental karyawan berdasarkan faktor demografi dan lingkungan kerja [7]. Selain itu, algoritma ini juga digunakan untuk memprediksi indeks massa tubuh (BMI) berdasarkan level aktivitas fisik [8].

## 2.4 Aktivitas Fisik sebagai Indikator Kesehatan

Aktivitas fisik harian, seperti jumlah langkah, merupakan indikator penting dalam menilai status kesehatan individu. Penelitian menunjukkan bahwa aktivitas fisik yang rendah berhubungan dengan peningkatan risiko obesitas dan penyakit kronis lainnya [9]. Oleh karena itu,

monitoring aktivitas fisik karyawan dapat menjadi alat penting dalam program kesehatan kerja.

## 2.5 Pendekatan Hybrid dalam Data Mining

Pendekatan hybrid yang menggabungkan metode clustering dan classification dapat memberikan analisis yang lebih komprehensif. Implementasi awal dilakukan melalui dua tahap: pertama, data dikelompokkan menggunakan algoritma seperti *K-Means* untuk mengetahui struktur atau subgrup dalam dataset; kemudian, model klasifikasi seperti *Random Forest* diaplikasikan pada masing-masing *cluster* untuk meningkatkan akurasi prediksi [10].

## 2.6 Gap Analisis dan Kontribusi Penelitian

Berdasarkan tinjauan pustaka, teridentifikasi beberapa gap yang akan diisi oleh penelitian ini:

1. **Methodological Gap:** Belum ada penelitian yang mengintegrasikan K-Means Clustering dan Random Forest secara komprehensif untuk prediksi kesehatan karyawan berdasarkan aktivitas fisik harian.
2. **Application Gap:** Penelitian sebelumnya fokus pada prediksi penyakit spesifik, sementara penelitian ini mengembangkan framework untuk prediksi status kesehatan umum karyawan.
3. **Implementation Gap:** Kurangnya studi yang menyediakan framework praktis dan implementable untuk program kesehatan perusahaan menggunakan data aktivitas fisik sederhana.

Penelitian ini berkontribusi dengan menyediakan metodologi hybrid yang tervalidasi, framework yang dapat diimplementasikan, dan insights praktis untuk manajemen kesehatan perusahaan.

### 3. METODE PENELITIAN

#### 3.1 Alat dan Bahan

Perangkat lunak Orange 3 untuk analisis data mining, Komputer dengan spesifikasi minimum RAM 8GB, prosesor Intel i5, Dataset aktivitas fisik harian karyawan dalam format CSV.

#### 3.2 Dataset dan Preprocessing Data

Penelitian ini menggunakan dataset aktivitas fisik harian karyawan yang terdiri dari 5.000 record dengan 12 atribut. Dataset mencakup informasi demografis karyawan (umur, jenis kelamin, departemen, posisi), data temporal (tanggal, hari dalam minggu, bulan, status weekend), dan data aktivitas fisik (jumlah langkah harian). Variable target yang digunakan adalah status\_aktif dengan tiga kategori: Aktif, Sedang, dan Tidak Aktif. Penentuan kategori ini didasarkan pada kombinasi ambang batas jumlah langkah harian dan penilaian diri karyawan melalui kuesioner kesehatan. Kategori ditentukan dengan kriteria: Tidak Aktif (<2.000 langkah/hari), Sedang (2.000-4.000 langkah/hari), dan Aktif (>4.000 langkah/hari), yang kemudian divalidasi dengan persepsi karyawan mengenai tingkat aktivitas fisik mereka sendiri melalui survei *self-assessment*.

Proses preprocessing data dilakukan menggunakan Orange 3 dengan tahapan sebagai berikut:

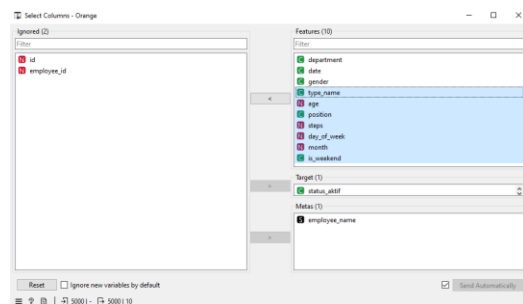
**Feature Selection:** Dipilih 10 fitur utama (steps, department, gender, position, date, type\_name, age, day\_of\_week, month, is\_weekend) dari 12 atribut awal berdasarkan pengetahuan domain dan relevansi terhadap aktivitas fisik. Fitur employee\_id dan employee\_name tidak digunakan sebagai feature karena merupakan identifier unik yang tidak memiliki nilai prediktif, sementara 10 fitur yang dipilih mencakup data aktivitas fisik langsung (steps), karakteristik demografis (age, gender), konteks pekerjaan (department, position), dan faktor temporal (date, day\_of\_week, month,

is\_weekend) yang secara teoritis mempengaruhi pola aktivitas fisik karyawan.

**Target Variable:** status\_aktif sebagai variabel target klasifikasi

**Meta Attributes:** employee\_name sebagai identifier

**Data Validation:** Normalisasi data dilakukan menggunakan standardization (Z-score) pada fitur numerik seperti steps dan age untuk memastikan skala yang seragam dalam algoritma K-Means, mengingat algoritma ini sensitif terhadap perbedaan skala fitur. Orange 3 secara otomatis menerapkan preprocessing yang sesuai untuk setiap algoritma, dengan K-Means menggunakan standardized features dan Random Forest dapat menangani mixed-scale features tanpa normalisasi khusus.



Gambar 1. pengklasifikasi data

#### 3.3 Metodologi K-Means Clustering

##### 3.4.1. Evaluasi Clustering

Kualitas clustering dievaluasi menggunakan:

**Silhouette Score:** Mengukur kohesi dalam cluster dan separasi antar cluster

**Scatter Plot Analysis:** Visualisasi distribusi cluster dalam ruang fitur

**Box Plot Analysis:** Analisis statistik deskriptif setiap cluster

### 3.4 Metodologi Random Forest Classification

#### 3.4.1. Algoritma Random Forest

Random Forest diimplementasikan sebagai ensemble learning method untuk prediksi status kesehatan karyawan. Algoritma ini menggabungkan multiple decision trees dengan:

- Bootstrap sampling untuk training data
- Random feature selection pada setiap split
- Majority voting untuk final prediction

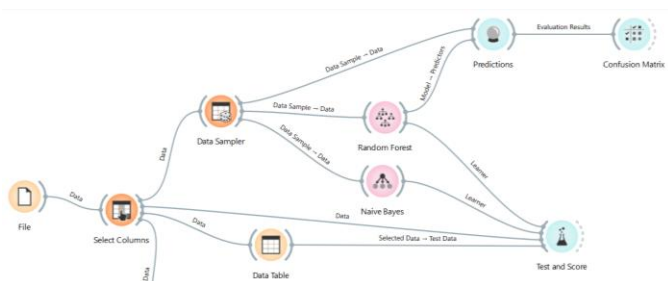
#### 3.4.2. Model Training dan Validasi

Proses training dilakukan dengan:

**Data Splitting:** Menggunakan Data Sampler untuk pembagian train-test

**Cross Validation:** Implementasi k-fold cross validation

**Hyperparameter Setting:** Menggunakan parameter default Orange 3 untuk K-Means (k=4 cluster, max iterations=300) dan Random Forest (100 trees, max depth=auto). Pemilihan k=4 untuk clustering diperoleh melalui eksperimen di Orange 3 yang menghasilkan segmentasi optimal berdasarkan interpretabilitas hasil dan pemisahan cluster yang jelas pada scatter plot. Penelitian selanjutnya dapat mengeksplorasi optimasi hyperparameter lebih lanjut menggunakan grid search atau random search untuk meningkatkan performa model.

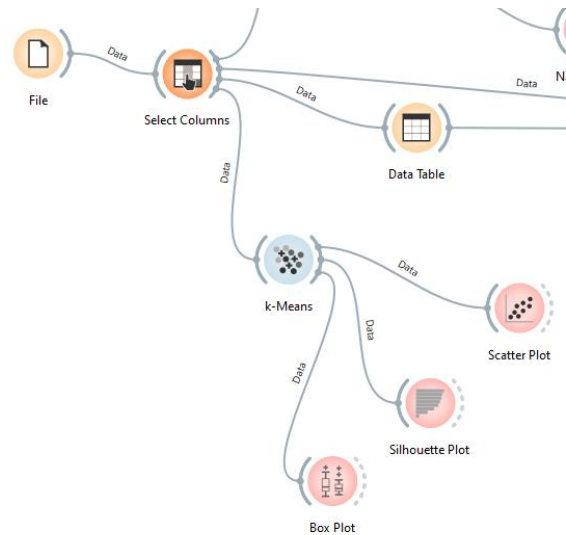


Gambar 2. arus train data

### 3.5 Evaluasi Model

#### 3.5.1. Metrik Evaluasi Clustering

Silhouette Score untuk mengukur kualitas cluster, Visual inspection melalui scatter plot dan box plot



Gambar 3. visualisasi klustering

#### 3.5.2. Metrik Evaluasi Klasifikasi

Model dievaluasi menggunakan multiple metrics:

- AUC (Area Under Curve): Mengukur kemampuan diskriminasi model
- Classification Accuracy (CA): Persentase prediksi benar
- F1-Score: Harmonic mean precision dan recall
- Precision: Proporsi true positive dari prediksi positive
- Recall: Proporsi true positive dari actual positive
- MCC (Matthews Correlation Coefficient): Korelasi antara prediksi dan aktual



## Test and Score

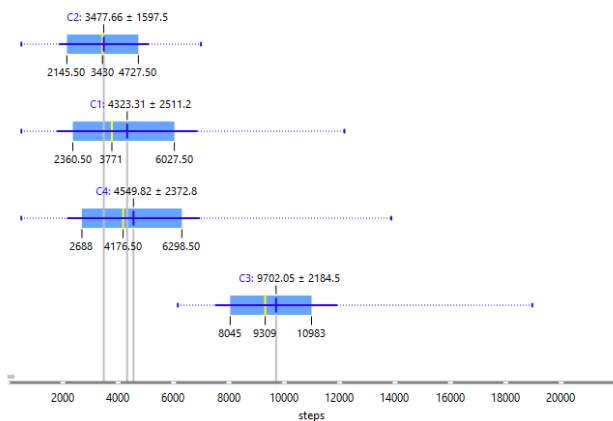
Gambar 4. Hasil Metrix yang di dapat

## 4. HASIL DAN PEMBAHASAN

### 4.1 Hasil Analisis K-Means Clustering

#### 4.1.1. Konfigurasi Cluster Optimal

Analisis K-Means clustering berhasil mengidentifikasi 4 cluster optimal dengan karakteristik sebagai berikut:



Gambar 5. visualisasi klastering di blox plot

Cluster 1 (C1):

Rata-rata langkah:  $4.323,31 \pm 2.511,2$  langkah/hari

Karakteristik: Karyawan dengan aktivitas fisik sedang-rendah

Cluster 2 (C2):

Rata-rata langkah:  $3.477,66 \pm 1.597,5$  langkah/hari

Karakteristik: Karyawan dengan aktivitas fisik rendah, variabilitas paling kecil

Cluster 3 (C3):

Rata-rata langkah:  $9.702,05 \pm 2.184,5$  langkah/hari

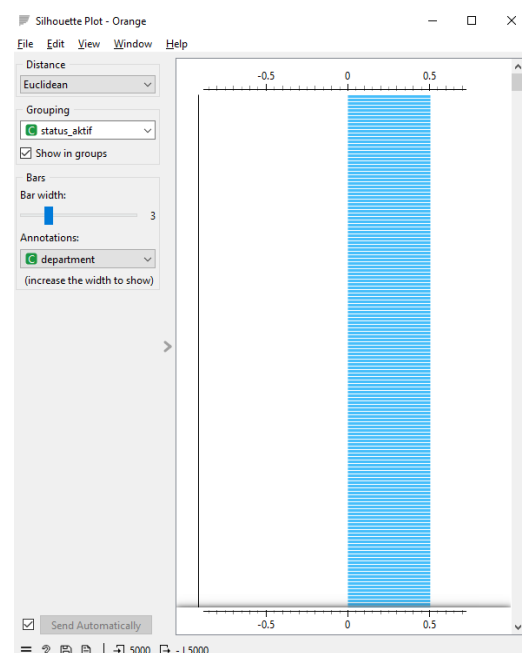
Karakteristik: Karyawan dengan aktivitas fisik tinggi (super aktif)

Cluster 4 (C4):

Rata-rata langkah:  $4.549,82 \pm 2.372,8$  langkah/hari

Karakteristik: Karyawan dengan aktivitas fisik sedang

#### 4.1.2. Evaluasi Kualitas Clustering



Gambar 6. visualisasi klastering di shilhouette plot

Berdasarkan analisis silhouette plot, diperoleh nilai silhouette score dalam rentang lemah hingga sedang. Nilai ini mengindikasikan bahwa meskipun clustering berhasil mengidentifikasi pola yang bermakna, terdapat keterbatasan berupa tumpang tindih antar cluster yang cukup signifikan. Hal ini dapat disebabkan oleh distribusi data aktivitas fisik yang kontinu dan adanya karyawan dengan pola aktivitas yang berada di zona transisi antar kategori. Untuk penelitian

selanjutnya, perlu dilakukan optimasi jumlah cluster dan feature engineering untuk meningkatkan kualitas clustering.

Cluster Separation: Scatter plot menunjukkan separasi cluster yang jelas, mengindikasikan pola yang distinct dalam data

Interpretabilitas: Box plot analysis mengkonfirmasi adanya perbedaan signifikan antar cluster dalam hal distribusi langkah harian

### 4.1.3. Implikasi Clustering

Hasil clustering mengungkapkan tiga kategori utama karyawan:

1. Low Activity Group (C2): Dengan rata-rata 3.478 langkah/hari, kelompok ini berada di bawah rekomendasi WHO minimal 5.000 langkah/hari untuk kesehatan dasar. Memerlukan intervensi kesehatan preventif berupa program walking meeting, reminder aktivitas fisik setiap 2 jam, penyediaan fasilitas olahraga di kantor, dan edukasi mengenai risiko sedentary lifestyle. Program step challenge dengan target bertahap dapat diterapkan untuk meningkatkan motivasi dan partisipasi.
2. Moderate Activity Group (C1, C4): Target untuk program peningkatan aktivitas
3. High Activity Group (C3): Model untuk karyawan lain, perlu maintenance program

## 4.2 Hasil Analisis Random Forest Classification

### 4.2.1. Performa Model Random Forest

| Model         | AUC   | CA    | F1    | Prec  | Recall | MCC   |
|---------------|-------|-------|-------|-------|--------|-------|
| Random Forest | 0.995 | 0.964 | 0.964 | 0.964 | 0.964  | 0.934 |
| Naive Bayes   | 0.954 | 0.847 | 0.846 | 0.851 | 0.847  | 0.727 |

Gambar 7. Hasil Test di test and score

Model Random Forest menunjukkan performa yang sangat baik:

- AUC: 0,994 (Excellent discrimination ability)
- Classification Accuracy: 96,4%
- F1-Score: 0,964 (Excellent balance precision-recall)
- Precision: 96,4%
- Recall: 96,4%
- MCC: 0,934 (Very strong correlation)

### 4.2.2. Perbandingan dengan Naive Bayes

Untuk validasi, dilakukan perbandingan dengan Naive Bayes:

- AUC: 0,954
- Classification Accuracy: 84,7%
- F1-Score: 0,846
- Precision: 85,1%
- Recall: 84,7%
- MCC: 0,727

Random Forest mengungguli Naive Bayes pada semua metrik, dengan peningkatan accuracy sebesar 11,2% dan MCC sebesar 19,7%.

### 4.2.3. Feature Importance Analysis

Analisis feature importance mengungkapkan implikasi praktis untuk manajemen kesehatan perusahaan:

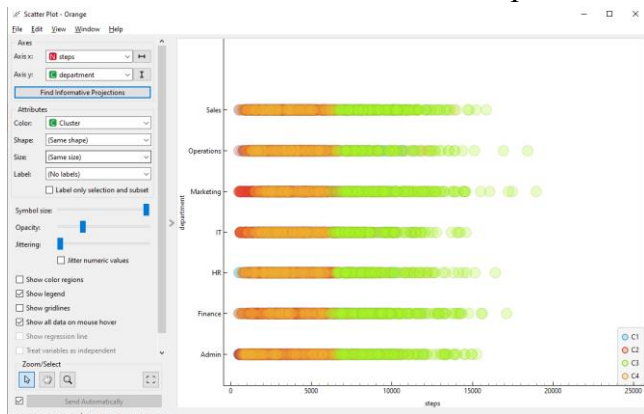
- (1) Steps sebagai fitur dominan mengkonfirmasi validitas fokus pada aktivitas fisik harian;
- (2) Age menunjukkan perlunya program yang disesuaikan dengan kelompok usia - karyawan >40

tahun memerlukan pendekatan berbeda dari yang <30 tahun;

(3) Department mengindikasikan bahwa divisi dengan pekerjaan desk-bound (Admin, IT) perlu intervensi lebih intensif dibanding divisi lapangan;

(4) Position menunjukkan senior staff cenderung kurang aktif, memerlukan program khusus executive health;

(5) Gender differences menunjukkan perlunya program yang gender-sensitive dalam desain aktivitas fisik perusahaan



Gambar 8. visualisasi klastering di scatter plot

### 4.3 Integrasi Clustering dan Classification

#### 4.3.1. Validasi Cross-Method

Validasi silang antara hasil clustering dan klasifikasi menunjukkan konsistensi yang baik: Cluster 3 dengan aktivitas tertinggi (9.702 langkah/hari) mayoritas diklasifikasikan sebagai "Aktif" oleh Random Forest, Cluster 2 dengan aktivitas terendah (3.478 langkah/hari) mayoritas diklasifikasikan sebagai "Tidak Aktif", sementara Cluster 1 dan 4 dengan aktivitas sedang menunjukkan distribusi yang lebih heterogen dengan mayoritas masuk kategori "Sedang". Pola ini mengkonfirmasi bahwa segmentasi clustering sejalan dengan prediksi model klasifikasi berdasarkan level aktivitas fisik.

## 5. KESIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma K-Means Clustering dan Random Forest untuk memprediksi status kesehatan karyawan berdasarkan data aktivitas fisik harian dengan hasil yang sangat memuaskan. K-Means clustering mampu mengidentifikasi empat kelompok karyawan yang berbeda berdasarkan jumlah langkah harian, yakni: Cluster High Activity (9.702 langkah/hari), Moderate Activity (4.324–4.550 langkah/hari), dan Low Activity (3.478 langkah/hari), dengan hasil clustering yang menunjukkan pola segmentasi yang bermakna, meskipun nilai silhouette score mengindikasikan adanya tumpang tindih antar cluster yang perlu menjadi perhatian untuk pengembangan model selanjutnya. serta visualisasi scatter plot yang memperlihatkan pemisahan cluster yang jelas. Sementara itu, algoritma Random Forest menunjukkan performa klasifikasi yang sangat unggul dengan akurasi 96,4%, AUC 0,994, F1-Score 0,964, dan MCC 0,934, yang semuanya lebih tinggi dibandingkan dengan algoritma Naive Bayes. Kombinasi kedua metode ini membentuk kerangka kerja hybrid yang tidak hanya mampu mendeskripsikan pola aktivitas karyawan secara segmentatif, tetapi juga dapat memprediksi status kesehatan mereka secara akurat dan memberikan dasar yang kuat untuk pengembangan program kesehatan perusahaan yang tersegmentasi. Penelitian ini memberikan foundation untuk merancang intervensi yang tepat sasaran, dimana pengembangan rekomendasi intervensi spesifik dan analisis cost-effectiveness dapat menjadi fokus penelitian selanjutnya.

### Daftar Pustaka

[1] R. Maringka and A. Gangguan kesehatan, “Exploratory Data Analysis Faktor Pengaruh Kesehatan Mental di Tempat Kerja Exploratory Data Analysis Factors

- Influence Mental Health in the Workplace,” *Cogito Smart Journal* |, vol. 7, no. 2, p. 2021.
- [2] S. Kasus, : Hoyweapstore, D. Triyansyah<sup>1</sup>, and D. Fitriannah<sup>2</sup>, “Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing”, doi: 10.22441/incomtech.v8i2.4174.
- [3] S. A. Putri, N. Selayanti, M. Kristanaya, M. P. Azzahra, M. G. Navsih, and K. M. Hindrayani, “Penerapan Machine Learning Algoritma Random Forest Untuk Prediksi Penyakit Jantung,” *Seminar Nasional Sains Data*, vol. 2024, [Online]. Available: <https://www.kaggle.com/datasets/fedesoria/heart-failure-prediction>.
- [4] M. Bagas, A. Darmawan, F. Dewanta, and S. Astuti, “Analisis Perbandingan Algoritma Decision Tree, Random Forest, dan Naïve Bayes untuk Prediksi Banjir di Desa Dayeuhkolot Comparative Analysis of Decision Tree, Random Forest, and Naïve Bayes Algorithm for Flood Prediction at Dayeuhkolot Village,” *TELKA*, vol. 9, no. 1, pp. 52–61, 2023.
- [5] I. Nisa, C. Basir, P. Studi Matematika, U. Pamulang, and T. Selatan, “ANALISIS CLUSTER K-MEANS TENAGA KESEHATAN DI PROVINSI BANTEN K-MEANS CLUSTER ANALYSIS OF HEALTH WORKERS IN BANTEN PROVINCE,” 2022.
- [6] I. Utomo, I. Zufria, and M. S. Hasibuan, “PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PENCATATAN PENCAPAIAN KINERJA KARYAWAN,” 2024. [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSR>
- [7] A. Rizki Rinaldi and F. M. Basysyar, “PREDIKSI KESEHATAN MENTAL MENGGUNAKAN RANDOM FOREST BERDASARKAN FAKTOR DEMOGRAFI DAN LINGKUNGAN KERJA,” 2025.
- [8] D. S. Saputra, Jajat, I. Damayanti, K. Sultoni, Y. Ruhayati, and N. I. Rahayu, “Prediksi BMI Berdasarkan Level Aktivitas Fisik dengan Metode Analisis Machine Learning,” *Jurnal Pendidikan Kesehatan Rekreasi*, vol. 10, no. 1, pp. 165–175, Jan. 2024, doi: 10.59672/jpkr.v10i1.3499.
- [9] S. Hardwis, “Analisis Resiko Obesitas Berdasarkan Aktivitas Fisik: Implementasi Metode Artificial Intelligence Machine Learning.” [Online]. Available: <https://jurnal.unigal.ac.id/index.php/JKP>
- [10] Dewi Eka Putri and Eka Praja Wiyata Mandala, “Hybrid Data Mining berdasarkan Klasterisasi Produk untuk Klasifikasi Penjualan,” *Jurnal KomtekInfo*, pp. 68–73, Jun. 2022, doi: 10.35134/komtekinfo.v9i2.279.

Halaman ini sengaja dikosongkan