

## DAFTAR PUSTAKA

- Adi Zulkarnaen Saputra, N. S. (2023). Klasterisasi Nilai Ujian Sekolah Menggunakan Metode Algoritma KMeans. *Jurnal Janitra Informatika dan Sistem Informasi*, 1-9.
- Darmawan, D. D. (2024). METODOLOGI PENELITIAN DALAM PENGEMBANGAN APLIKASI STREAMLIT: STUDI KASUS VISUALISASI DATA GEMPA BUMI DI INDONESIA. *researchgate*.
- F N Dhewayani, D. A. (2022). Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi dan Informasi*, 2088-2270.
- Hidayat, R. &. (2020). Penerapan Metode Data Mining untuk Memprediksi Penyakit Jantung Menggunakan Algoritma K-Means. *Jurnal Teknologi dan Sistem Komputer*, 1-8.
- Khoirudin. (2024). Klasifikasi Penyakit Jantung Menggunakan Perbandingan Algoritma Decision Tree. *Jurnal Ilmiah Manajemen Informatika dan Komputer*, 19-25.
- Nazar, R. (2024). IMPLEMENTASI PEMROGRAMAN PYTHON MENGGUNAKAN GOOGLE COLAB. *Jurnal Informatika dan Komputer*, 2089 – 4384.
- Nurul Maulida Surbakti, A. ,.-A. (2024). Penggunaan Bahasa Pemrograman Python dalam Pembelajaran Kalkulus Fungsi Dua Variabel. *Jurnal Matematika, Ilmu pengetahuan Alam, Kebumihan dan Angkasa*, 98-107.
- Organization, W. H. (2021, Juni 11). *Penyakit kardiovaskular (PKV)*. Retrieved from World Health Organization: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- Pramudito, A. &. (2021). Klasterisasi Data Pelanggan Menggunakan Algoritma K-Means untuk Meningkatkan Strategi Pemasaran. *Jurnal Ilmiah Teknologi Informasi*, 45-52.

Sari, R. F. (2022). Implementasi Data Mining untuk Analisis Sentimen pada Ulasan Produk Menggunakan Algoritma Naïve Bayes. *Jurnal Sistem Informasi*, 23-30.

Setiawan, A. &. (2023). Penerapan Data Mining untuk Memprediksi Penyakit Diabetes Menggunakan Algoritma Klasifikasi. *Jurnal Informatika*, 15-22.

Sitompul, I. S. (2024). Analisis Pelaksanaan Sistem Informasi Kesehatan Dinas Kesehatan Provinsi Riau. *Journal of Healthcare Technology and Medicine*.

Yunita, F. (2020). PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING. *SISTEMASI*.

Yusriana Chusna Fadilah, A. S. (2025). APPLYING K-MEANS CLUSTERING FOR GROUPING PAPUA'S DISTRICTS BASED ON POVERTY INDICATORS ANALYSIS. *Jurnal Ilmu Pengetahuan dan Teknologi Komputer*, 2685-8223.



## LAMPIRAN

### Lampiran 1. Source Code

```
# Impor Library yang Dibutuhkan

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans
```

```
# Memasukan Dataset
df = pd.read_csv('/content/heart.csv')
df.head()
```

```
# Tampilkan informasi dataset
print("Info Dataset:")
print(df.info())
```

```
# Preproses Data
# Mengisi missing values jika ada
df = df.dropna() # Contoh sederhana: menghapus baris dengan missing values
```

```
# Check for missing values
print("\nMissing values in each column:")
print(df.isnull().sum())
```

```
# Pilih atribut yang akan digunakan untuk clustering
selected_features = ['age', 'sex', 'cp', 'trestbps', 'chol', 'thalach',
                    'oldpeak', 'slope', 'ca']
data = df[selected_features]
data
```

```
# Normalisasi data
scaler = StandardScaler()
scaled_data = scaler.fit_transform(data)
scaled_data = np.round(scaled_data, 2)
scaled_data
```

```

k = 2
data = np.random.rand(10, 9)

# Centroid awal
initial_centroids = np.array([
    [0.83, -1.51, -0.92, -0.43, -0.72, 0.6, -0.91, 1, -0.73],
    [-0.16, -1.51, 1.03, -0.21, -0.58, -1.48, -0.91, 1, -0.73]
])
print("\nCentroid Awal (Random Initialization):")
print(initial_centroids)

```

```

def euclidean_distance(x1, x2):
    return np.sqrt(np.sum((x1 - x2) ** 2))

def calculate_distances(data, centroids):
    distances = []
    for point in data:
        dist_to_centroids = [euclidean_distance(point, centroid) for
centroid in centroids]
        distances.append(dist_to_centroids)
    return np.array(distances)

distances = calculate_distances(scaled_data, initial_centroids)

labels = np.argmin(distances, axis=1)

distance_table = []
for idx, row in enumerate(distances):
    closest_centroid_distance = row[labels[idx]]
    cluster_label = labels[idx]
    distance_table.append([idx, scaled_data[idx], closest_centroid_distance,
cluster_label])

distance_df = pd.DataFrame(distance_table, columns=["Sample", "Data",
"Jarak", "Label"])

print("\nTabel Jarak dan Label Cluster:")
distance_df.head(10)

```

```

import matplotlib.pyplot as plt
from sklearn.cluster import KMeans

k_range = range(1, 11)
inertia = []

for k in k_range:
    kmeans = KMeans(n_clusters=k, random_state=0)
    kmeans.fit(data)
    inertia.append(kmeans.inertia_)

plt.figure(figsize=(8, 5))
plt.plot(k_range, inertia, marker='o')
plt.title('Metode Elbow untuk Menentukan Jumlah Cluster Optimal')
plt.xlabel('Jumlah Cluster (k)')
plt.ylabel('Inertia')
plt.xticks(k_range)
plt.grid(True)
plt.show()

```

```

# Terapkan K-Means Clustering dengan 2 Cluster
optimal_k = 2
kmeans = KMeans(n_clusters=optimal_k, random_state=42)
df['cluster'] = kmeans.fit_predict(scaled_data)

```

```

# Inisialisasi model K-Means dengan jumlah klaster = 2
kmeans = KMeans(n_clusters=2, random_state=42)

# Fit model ke data yang sudah discaling (scaled_data) dan memiliki 1025 baris
# scaled_data = scaler.fit_transform(data_for_clustering) # Pastikan data_for_clustering adalah data yang ingin Anda cluster, dengan 1025 baris
kmeans.fit(scaled_data) # Ganti X_scaled dengan scaled_data

# Tambahkan label klaster ke dataset
df['Cluster'] = kmeans.labels_

# Tampilkan centroid
centroids = kmeans.cluster_centers_
print("Centroid Klaster:\n", centroids)

# Tampilkan hasil klasterisasi
print(df[['age', 'Cluster']].head(10))

```

```
# Centroid dan Label
centroids = kmeans.cluster_centers_

# Buat DataFrame untuk centroid
centroid_df = pd.DataFrame(centroids, columns=selected_features)

# Tambahkan label cluster sebagai indeks
centroid_df.index = [f'Cluster {i}' for i in range(len(centroids))]

# Output DataFrame centroid
print("Centroid DataFrame:")
centroid_df
df
```

```
# Ganti label cluster untuk mempermudah interpretasi
# Anggap Cluster 0 adalah "Tidak Akut" dan Cluster 1 adalah "Akut"
df['cluster'] = df['cluster'].replace({0: 'Tidak Akut', 1: 'Akut'})
df
```

```
# Visualisasi Cluster
plt.figure(figsize=(10, 6))
sns.scatterplot(data=df, x="age", y="thalach", hue="cluster",
palette="Set1", s=100)
plt.scatter(centroids[:, 0], centroids[:, 2], color='blue', marker='X',
s=200, label='Centroids')
plt.title("K-Means Clustering (age vs thalach)")
plt.xlabel("Age")
plt.ylabel("Thalach")
plt.legend()
plt.show()
```

```
# Hitung Silhouette Score
sil_score = silhouette_score(scaled_data, df['cluster'])
# Menampilkan Silhouette Score
print(f"Silhouette Score untuk K-Means Clustering dengan 2 Cluster:
{sil_score:.4f}")
```

```
# Hitung rata-rata usia untuk setiap kombinasi cluster dan jenis kelamin
average_age_per_cluster_gender = df.groupby(['cluster',
'sex'])['age'].mean().reset_index()
```

```
# Visualisasi dengan bar plot
plt.figure(figsize=(12, 6))
sns.barplot(x='cluster', y='age', hue='sex',
data=average_age_per_cluster_gender, palette="Set2")
plt.title('Rata-rata Usia Berdasarkan Cluster dan Jenis Kelamin')
plt.xlabel('cluster')
plt.ylabel('Rata-rata Usia')
plt.show()
```

```
# Visualisasi hasil klasterisasi
import matplotlib.pyplot as plt

# Plot hasil clustering menggunakan dua dimensi (contoh: age dan chol)
plt.scatter(scaled_data[:, 0], scaled_data[:, 4], c=kmeans.labels_,
cmap='viridis', s=50) # Replace X_scaled with scaled_data
plt.scatter(centroids[:, 0], centroids[:, 4], c='red', marker='x', s=200,
label='Centroids')
plt.title('Hasil Klasterisasi K-Means')
plt.xlabel('Age (Scaled)')
plt.ylabel('Chol (Scaled)')
plt.legend()
plt.show()
```

```
# Analisis dan Simpan Data dengan Cluster
# Tampilkan jumlah data pada masing-masing cluster
cluster_counts = df['cluster'].value_counts()
print("\nJumlah Data di Setiap Cluster:")
print(cluster_counts)
```

```
# Visualisasi cluster
plt.figure(figsize=(8, 6))
sns.scatterplot(data=df, x='age', y='thalach', hue='cluster',
palette='Set1', s=100)
plt.title('Hasil Clustering (2 Cluster) - Age vs Thalach')
plt.xlabel('Age')
plt.ylabel('Thalach')
plt.legend()
plt.show()
```

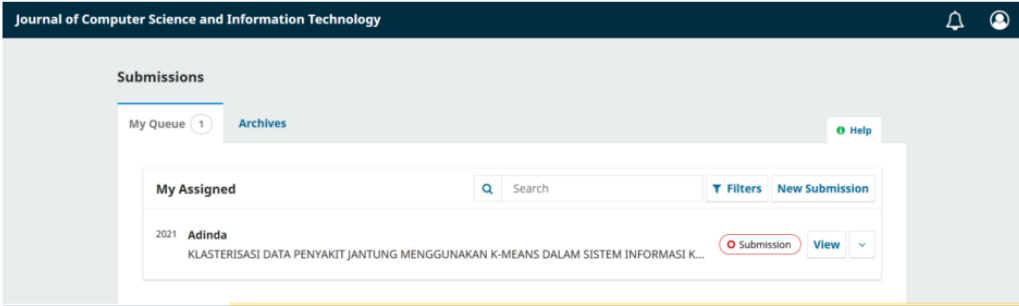
## Lampiran 2. Turnitin

### KLASTERISASI DATA PENYAKIT JANTUNG MENGGUNAKAN K-MEANS DALAM SISTEM INFORMASI KESEHATAN.docx

#### ORIGINALITY REPORT

|                 | 4%                                                                                                                                                                    | 4%               | 2%           | 2%             |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|--------------|----------------|
|                 | SIMILARITY INDEX                                                                                                                                                      | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |
| PRIMARY SOURCES |                                                                                                                                                                       |                  |              |                |
| 1               | doczz.net                                                                                                                                                             | Internet Source  | <1 %         |                |
| 2               | repository.unfari.ac.id                                                                                                                                               | Internet Source  | <1 %         |                |
| 3               | repository.ub.ac.id                                                                                                                                                   | Internet Source  | <1 %         |                |
| 4               | 123dok.com                                                                                                                                                            | Internet Source  | <1 %         |                |
| 5               | text-id.123dok.com                                                                                                                                                    | Internet Source  | <1 %         |                |
| 6               | Submitted to STIE Perbanas Surabaya                                                                                                                                   | Student Paper    | <1 %         |                |
| 7               | docplayer.info                                                                                                                                                        | Internet Source  | <1 %         |                |
| 8               | Submitted to University of North Carolina, Greensboro                                                                                                                 | Student Paper    | <1 %         |                |
| 9               | repository.upi.edu                                                                                                                                                    | Internet Source  | <1 %         |                |
| 10              | Submitted to Cardiff University                                                                                                                                       | Student Paper    | <1 %         |                |
| 11              | repository.fe.unj.ac.id                                                                                                                                               | Internet Source  | <1 %         |                |
| 12              | Elsa Safitri, Yurita Mailintina, Ria Efkelin.<br>"Gambaran Penggunaan Sistem Informasi Kesehatan Elektronik di Puskesmas Kalideres", Malahayati Nursing Journal, 2024 | Publication      | <1 %         |                |

### Lampiran 3. Bukti Submission & Letter of Acceptance



Journal of Computer Science and Information Technology

Submissions

My Queue 1 Archives Help

My Assigned

Search Filters New Submission

2021 Adinda

KLASTERISASI DATA PENYAKIT JANTUNG MENGGUNAKAN K-MEANS DALAM SISTEM INFORMASI K...

Submission View


Journal of Computer Science and Information Technology

← Back to Submissions

2021 / Adinda / KLASTERISASI DATA PENYAKIT JANTUNG MENGGUNAKAN K-MEANS DALAM SISTEM INFORMASI KESEHATAN Library

Workflow Publication

Submission Review Copyediting Production

Submission Files

14485 Draft Jurnal - Saskia Adinda - JCSIT.docx

February 20, 2025 Article Text

Download All Files

**Pusat Pengembangan dan Penerapan Hasil Penelitian dan Pengabdian Kepada Masyarakat (P3HPPKM)**

**JCSIT** Journal of Computer Science and Information Technology (JCSIT)

Alamat: Perum taman Asri 1 Blok C2 RT 31 RW 06 Palembang Sumatera Selatan 30149

**LETTER OF ACCEPTANCE (LOA)**

**Nomor: 65 / YNIM / PPPMI / JCSIT / 1 / 2025**

Berdasarkan hasil telaah tim Redaksi **Journal of Computer Science and Information Technology (JCSIT)**, menerangkan bahwa artikel yang dimuat atas nama author dibawah ini:

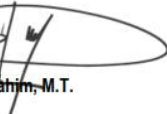
**Nama** : Saskia Adinda, Andrianingsih Andrianingsih

**Judul Artikel** : **KLASTERISASI DATA PENYAKIT JANTUNG MENGGUNAKAN K-MEANS DALAM SISTEM INFORMASI KESEHATAN**

**E-ISSN** : 3031-8467

Sudah diterima dan akan diterbitkan pada edisi Vol.2 No. 2, Bulan Maret 2025. Demikianlah *Letter of Accepted* (LOA) ini dibuat semoga dapat dipergunakan sebagaimana mustinya.

Palembang, 28 Febuari 2025  
Pimpinan Redaksi



Dr. Ali Ibrahim, M.T.



#### Lampiran 4. Dokumentasi Bimbingan

