I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694

CLUSTERING DENGAN SIMULATED ANNEALING UNTUK PENGELOMPOKAN TINDAKAN MENYAKITI DIRI SENDIRI PADA MAHASISWA SEMESTER AKHIR

Novansa Kurniatama Sumarna[1], Febri Ainun Jariyah[2], Ridwansyah[3*], Yamin Nuryamin[4]
[1.2,3,4] Universitas Bina Sarana Informatika
[1,2,3,4] Jl. Kramat Raya No. 98, Senen,
Jakarta Pusat 10450.
Email : novan1152@gmail.com

ABSTRAK

Masalah kesehatan mental, termasuk tindakan self-harm, semakin menjadi perhatian serius di kalangan mahasiswa, terutama pada mahasiswa semester akhir yang menghadapi tekanan akademik dan emosional yang tinggi. Penelitian ini bertujuan untuk mengimplementasikan metode K-Means Clustering yang dioptimalkan dengan Simulated Annealing untuk mengelompokkan mahasiswa berdasarkan pola tindakan self-harm. Proses penelitian melibatkan pengumpulan data melalui survei, preprocessing data, implementasi algoritma K-Means dengan optimasi Simulated Annealing, dan evaluasi hasil clustering menggunakan metrik seperti Silhouette Score, Davies-Bouldin Index, dan Calinski-Harabasz Index. Hasil penelitian menunjukkan bahwa metode K-Means yang dioptimalkan dengan Simulated Annealing mampu menghasilkan pengelompokan yang lebih efektif dan akurat dibandingkan K-Means standar, dengan kluster-kluster yang terbentuk memberikan gambaran yang jelas tentang kelompok mahasiswa berdasarkan tingkat stres, dukungan sosial, dan frekuensi self-harm. Penelitian ini diharapkan dapat memberikan kontribusi nyata dalam mendukung upaya deteksi dini dan pencegahan tindakan self-harm pada mahasiswa semester akhir, serta menjadi dasar bagi pengembangan sistem pendukung keputusan dalam pengelolaan kesehatan mental di lingkungan kampus.

Kata Kunci *: K-Means, Machine Learning, Mahasiswa, Simulated Annealing.*

ABSTRACT

Mental health problems, including acts of self-harm, are increasingly becoming a serious concern among students, especially in final semester students who face high academic and emotional pressure. This research aims to implement the K-Means Clustering method optimized with Simulated Annealing to group students based on self-harm action patterns. The research process involves collecting data through surveys, data preprocessing, implementing the K-Means algorithm with Simulated Annealing optimization, and evaluating clustering results using metrics such as Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index. The research results show that the K-Means method optimized with Simulated Annealing is able to produce groupings that are more effective and accurate than standard K-Means, with the clusters formed providing a clear picture of student groups based on stress levels, social support, and frequency of self- harm. It is hoped that this research can make a real contribution in supporting efforts for early detection and prevention of self-harm in final semester students, as well as being the basis for developing a decision support system in managing mental health in the campus environment.

Keywords: *K-Means, Machine Learning, Simulated Annealing, Student.*

1. PENDAHULUAN

Tindakan melukai diri sendiri atau *self-harm* merupakan salah satu bentuk gangguan psikologis yang kian marak terjadi, terutama pada kalangan mahasiswa tingkat akhir. Adanya tanggung jawab dan tuntutan kehidupan akademik pada mahasiswa tentu dapat menjadi bagian dari stress yang dialami oleh mahasiswa (Oktariani, Putri, & Sofah, 2021). Mahasiswa tingkat akhir berada dalam fase kritis yang menuntut penyelesaian skripsi, penentuan arah

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694

karier, dan adaptasi terhadap ekspektasi sosial (Asmiyah, Madjid, & Safriyani, 2025). Tekanan yang tidak ditangani dengan baik dapat berujung pada stres berkepanjangan dan perilaku menyakiti diri sebagai bentuk pelarian emosional (Hafid & Usop, 2025).

Tindakan self-harm sering tidak terdeteksi secara langsung karena dilakukan secara tertutup dan jarang dibagikan kepada orang lain (R. Yang et al., 2021). Kondisi ini membuat upaya pencegahan dan penanganan menjadi lebih sulit (Tarigan & Apsari, 2024). Oleh karena itu, diperlukan pendekatan berbasis data yang dapat membantu mengelompokkan individu-individu yang memiliki kecenderungan melakukan tindakan *self-harm* berdasarkan pola-pola psikologis yang teridentifikasi, seperti tingkat depresi, kecemasan, dan stres yang dapat diukur melalui kuesioner psikologis seperti DASS (Hou et al., 2025).

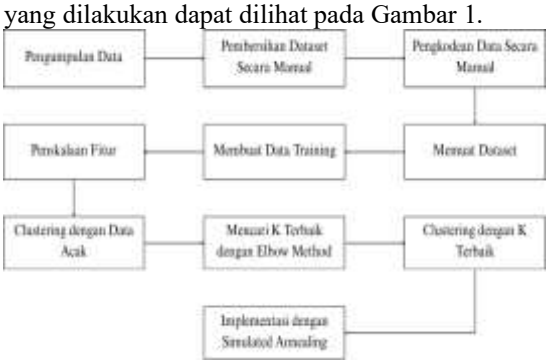
Salah satu metode yang dapat digunakan untuk mengelompokkan data tersebut adalah K-Means Clustering, yang efektif dalam membagi data ke dalam kelompok berdasarkan kemiripan nilai atribut (Suhardjono, Sugiarto, Yuliandari, Sudradjat, & Rohimah, 2023). Namun, metode ini memiliki kelemahan dalam pemilihan *centroid* awal, yang dapat menyebabkan hasil klaster kurang optimal atau terjebak pada solusi lokal (J. Yang, Wang, Yao, & Lin, 2021). Oleh karena itu, perlu dilakukan optimasi terhadap proses *clustering* untuk memperoleh hasil yang lebih akurat dan representatif (Ridwansyah, Riyanto, Hamid, Rahayu, & Purnama, 2022).

Untuk mengatasi keterbatasan tersebut, algoritma *Simulated Annealing* digunakan sebagai metode optimasi (H. Yang, Wen, & Geng, 2024). *Simulated Annealing* merupakan algoritma metaheuristik yang mampu melakukan pencarian global dan menghindari solusi lokal, sehingga dapat membantu dalam menemukan titik pusat klaster (*centroid*) yang lebih optimal (Putra, 2025). Dengan menggabungkan K-Means dan *Simulated Annealing*, diharapkan proses pengelompokan data mahasiswa berdasarkan kecenderungan *self-harm* dapat dilakukan dengan hasil yang lebih baik dan reliabel (Shaikh, Dong, Zheng, Wang, & Lin, 2024).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan metode pengelompokan data mahasiswa berdasarkan kecenderungan *self-harm* menggunakan kombinasi algoritma K-Means dan Simulated Annealing. Hasil pengelompokan ini sangat berguna dalam konteks pencegahan dan intervensi dini terhadap perilaku *self-harm*. Dengan mengetahui kelompok mahasiswa yang memiliki tingkat risiko tinggi, pihak kampus dapat menyusun program pendampingan dan konseling yang lebih tepat sasaran. Hal ini juga mendukung upaya peningkatan layanan kesehatan mental di lingkungan pendidikan tinggi, yang selama ini masih sering terabaikan.

2. METODE PENELITIAN

Dalam menyusun penelitian "Clustering dengan Simulated Annealing untuk Pengelompokan Tindakan Menyakiti Diri Sendiri pada Mahasiswa Semester Akhir" ini, terdapat proses penelitian yang berisi: kerangka penelitian, sumber data, teknik pengambilan data. Berikut adalah tahapan penelitian yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Langkah Penelitian

2.1 Pengumpulan Data

Pengumpulan data merupakan tahap penting dalam penelitian ini, karena kualitas data yang digunakan akan sangat memengaruhi hasil pengelompokan tindakan *self-harm* pada mahasiswa semester akhir. Data ini diperoleh melalui kuesioner yang dirancang khusus untuk mengidentifikasi pola perilaku *self-harm* dan faktor-faktor pemicunya Kuesioner terdiri dari pertanyaan yang terstruktur dan terstandar, sehingga mempermudah proses pengumpulan data. Data yang peneliti dapatkan berisi data kualitatif dan kuantitatif yang seluruhnya berjumlah 250 data. Berikut adalah sample data yang di dapatkan.

dalam algoritma. Klaster hasil optimasi terlihat lebih terpisah dan padat, menandakan peningkatan kualitas pengelompokan dibandingkan K-Means murni. Dengan optimasi ini, model berhasil memperjelas batas antar-klaster dan menghasilkan pengelompokan yang lebih reliabel.

Sebagai tahap lanjutan, jumlah klaster ditingkatkan menjadi K=4 untuk melihat apakah pengelompokan dapat memberikan hasil yang lebih detail.

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694

Gambar 11. Visualisasi Cluster Data dan Centroid

Pada Gambar 11 dapat terlihat kriteria berikut:

a. Klaster 0 (biru): tingkat rendah atau sehat
b. Klaster 1 (ungu): tingkat rendah–menengah
c. Klaster 2 (hijau): tingkat menengah

Gambar 2. Sample Data Awal

2.2. *Preprocessing Data*

Peneliti mengumpulkan 250 data melalui kuesioner Google Forms. Langkah pertama dalam preprocessing adalah **pembersihan data** (data cleaning), di mana data yang tidak valid, seperti data kosong, duplikat, atau outlier, dihapus. Proses ini memastikan bahwa data yang digunakan relevan dan berkualitas. Pembersihan data secara manual yang peneliti lakukan menghasilkan data bersih berjumlah 246 data.

2.3 *Pengkodean Data Secara Manual*

Setelah membersihkan data dan menghasilkan 246 data yang valid, peneliti selanjutnya melakukan pengkodean pada variabel jenis kelamin. Jenis kelamin responden dikodekan menjadi "JK". Selain itu, peneliti juga melakukan pengkodean terhadap tingkat self-harm yang dialami oleh responden. Variabel ini diberi kode "TS" agar dapat diintegrasikan dengan mudah dalam analisis data. Dimana perempuan diberi kode 0 dan laki-laki diberi kode 1. Tujuan dari pengkodean ini adalah untuk memudahkan pengelolaan dan analisis data lebih lanjut.

Dengan mengubah data jenis kelamin menjadi format kode numerik, peneliti dapat dengan mudah memasukkan dan memproses variabel ini ke dalam analisis statistik. Pengkodean data jenis kelamin menjadi JK, dengan 0 untuk perempuan dan 1 untuk laki-laki. Selain itu, peneliti juga melakukan pengkodean terhadap tingkat self-harm yang dialami oleh responden. Variabel ini diberi kode "TS". Pengkodean data jenis kelamin menjadi JK, dengan 0 untuk perempuan dan 1 untuk laki-laki, serta pengkodean tingkat self-harm menjadi TS dengan skala 1 hingga 4, merupakan langkah penting dalam mempersiapkan data untuk analisis yang lebih komprehensif nantinya di *Google Colab*

2.4 *Memuat Dataset*

Memuat dataset sederhananya ialah mengupload dataset. Dalam proses pemuatan dataset, penulis menggunakan *google colab* sebagai media untuk menjalankan *syntax Phyton* sebagai platform untuk menjalankan model *K-Means*.

2.5 Membuat Data *Training*

Setelah memuat dataset yang telah dibersihkan dan dikodekan ke dalam *Google Colab*, langkah selanjutnya adalah membuat data *training* untuk keperluan *clustering*. Data training diambil dari kolom 'KODE TS' dan 'KODE TS' pada dataset df. Penggunaan df[['KODE TS','KODE TS']].values akan mengekstrak nilai-nilai dari kedua kolom tersebut dan menyimpannya dalam variabel. Proses ini mempersiapkan data training yang akan digunakan untuk clustering. Dengan memiliki data training yang telah siap, peneliti dapat melanjutkan ke tahap selanjutnya dalam proses clustering data self-*harm*. Proses ini mempersiapkan data *training* yang akan digunakan untuk *clustering*. Dengan memiliki data *training* yang telah siap, peneliti dapat melanjutkan ke tahap selanjutnya dalam proses *clustering* data *self-harm*.

2.6 Penskalaan Fitur

Dalam proses penskalaan fitur atau *Feature Scaling* peneliti menggunakan metode *MinMaxScaler* dari *library sklearn.preprocessing* untuk melakukan penskalaan fitur. Pertama-tama peneliti mengimpor *MinMaxScaler* dari *sklearn.preprocessing*. Kemudian, peneliti membuat objek *scaler* dari kelas *MinMaxScaler*(). Selanjutnya, *scaler.fit_transform*(*x_train*) digunakan untuk menghitung parameter skala (minimum dan maksimum) berdasarkan data *training x_train*, lalu mentransformasikan *x_train* agar berada dalam rentang 0 hingga 1.

2.7 *Clustering* dengan Data Acak

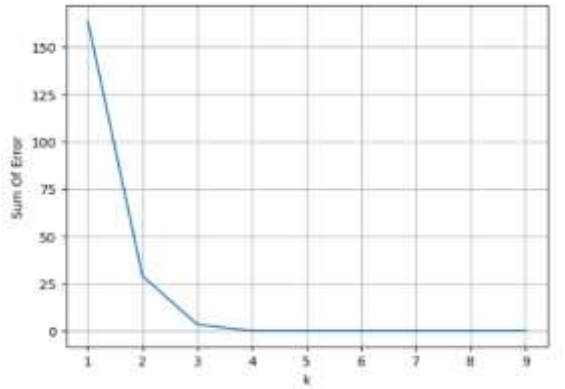
Dalam *clustering* menggunakan algoritma *K-Means*, peneliti akan mengelompokkan data

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694

berdasarkan tingkat *self-harm* yang dialami oleh responden. pertama-tama peneliti mengimpor K-Means dari sklearn.cluster. Kemudian, peneliti membuat objek *K-Means* dari kelas *K-Means*() dan menentukan jumlah *cluster* yang akan dibuat, yaitu 2 *cluster*. Hasil dari proses *clustering* ini disimpan dalam variabel y_*cluster*, yang berisi label *cluster* untuk setiap data dalam x_*train*. Contoh y_cluster: [1, 0]. Selanjutnya, membuat kolom baru yang bernama "Kode Hasil (HSL)" yang diisi dengan data dari y_*cluster* dan simpan di dalam *dataframe*. Kolom baru ini dimaksudkan sebagai data yang berisi hasil pertama dari *clustering*. Berikut adalah *sample* data dari KODE HSL: [0, 1]. Setelah KODE HSL terbentuk maka diganti nilai dari KODE HSL. Nilai 0 diganti menjadi "SEHAT" sedangkan nilai 1 diganti menjadi "*SELF-HARM*". Hasil dari penggantian nilai-nilai pada KODE HSL menjadikan data lebih mudah dimengerti dan dipahami.

2.8 *Clusthering* dengan K Terbaik

Menjalankan algoritma *K-Means* dengan metode *Elbow-Method* untuk mencari K Terbaik. K terbaik disini dimaksudkan untuk jumlah *cluster* yang paling optimal, sehingga hasil dari *clustering* selanjutnya akan menggunakan nilai K jumlah cluster yang ditunjukan dari hasil *Elbow Method*. Metode *Elbow* menentukan jumlah *cluster* yang paling tepat berdasarkan perubahan nilai *Sum of Squared Errors (SSE*) atau inertia yang ditunjukan dengan grafik paling patah atau siku (*elbow*).



Gambar 3. Grafik Elbow Method

2.9 *Mencari K Terbaik dengan Elbow Method*

Setelah melihat grafik yang dihasilkan akan menunjukkan perubahan nilai *inertia* berdasarkan jumlah *cluster* yang berbeda. Titik 3 di mana grafik menunjukkan perubahan tajam atau "siku" (*elbow*) menunjukkan jumlah *cluster* yang paling optimal yaitu 3 *cluster*. 3 *cluster* ini selanjutnya digunakan untuk *clustering* K-Means sehingga menghasilkan data *cluster* seperti *sample* 0, 1 dan 2 dengan *centroid* 3 *cluster*. Setelah menentukan jumlah *cluster* yang optimal menggunakan metode *Elbow*, maka akan membuat dataset baru yang mencakup label *cluster* untuk setiap data.

2.10 Implementasi dengan *Simulated Annealing*

Dalam konteks *K-Means*, *simulated annealing* dapat digunakan untuk mencari nilai *centroid* yang optimal, sehingga proses *clustering* dapat menghasilkan kelompok-kelompok yang lebih terdefinisi dengan baik. Dengan mengombinasikan *K-Means* dan *Simulated Annealing*, peneliti dapat memperoleh hasil *clustering* yang optimal dan lebih akurat dibandingkan hanya menggunakan *K-Means* biasa. Dengan demikian, partikel berupaya untuk bergerak menuju solusi optimal dengan menggabungkan pengalaman individu dan informasi kolektif. Setelah beberapa iterasi, partikel akan menemukan posisi *centroid* yang lebih baik, menghasilkan kluster yang lebih stabil dan representatif, serta meningkatkan kualitas pengelompokan.

3. HASIL DAN PEMBAHASAN

Hasil pengujian dalam penelitian ini dianalisis menggunakan **Google Colab**, dengan tujuan untuk mengevaluasi kinerja algoritma **K-Means Clustering**, **Elbow Method**, dan **Simulated Annealing** dalam proses pengelompokan data mahasiswa berdasarkan tingkat kecenderungan *self-harm*.

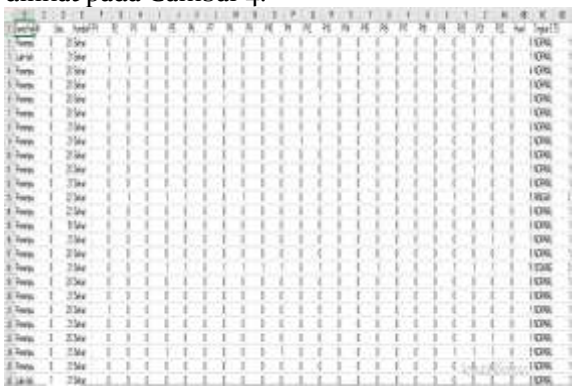
Proses analisis diawali dengan penerapan **Elbow Method** untuk menentukan jumlah klaster optimal yang merepresentasikan variasi pola psikologis mahasiswa. Setelah jumlah klaster ditetapkan, algoritma **K-Means** digunakan untuk melakukan pengelompokan awal berdasarkan nilai-nilai pada variabel depresi, kecemasan, dan stres. Selanjutnya, hasil klaster tersebut dioptimasi menggunakan **Simulated Annealing** untuk memperoleh posisi centroid yang lebih stabil dan menghindari solusi lokal, sehingga kualitas klaster menjadi lebih akurat dan reliabel.

3.1 Data

Data yang digunakan dalam penelitian ini merupakan hasil pengumpulan kuesioner terhadap 246 mahasiswa semester akhir yang bertujuan untuk mengidentifikasi kecenderungan tindakan menyakiti diri sendiri (*self-harm*). Variabel utama yang dianalisis adalah jenis kelamin (JK) dan tingkat *self-harm* (TS) yang dikodekan dalam bentuk numerik, di mana 0 merepresentasikan perempuan, 1 untuk laki-laki, dan tingkat *self-harm* dinilai dengan skala 1 hingga 4. Proses analisis dilakukan menggunakan *Google Colab* dengan algoritma *K-Means Clustering* dan dioptimasi menggunakan *Simulated Annealing* (SA).

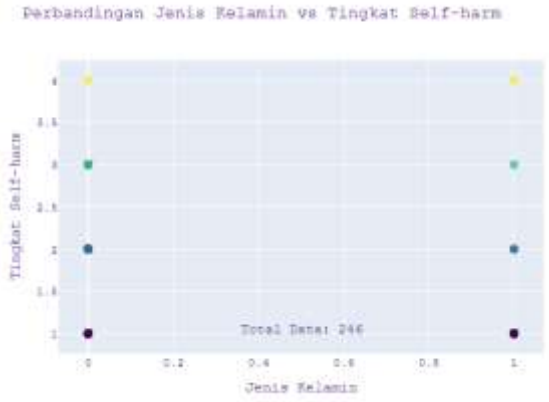
INFORMATIKA
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694

Data yang sudah di kumpulkan dan sudah melalui proses *preprocessing* data, pengkodean data, pengolahan data, dan sudah dilakukan penskalaan fitur dengan hasil data siap dipakai yang dapat dilihat pada Gambar 4.



Gambar 4. Data Siap Olah

Pada Gambar 4 menerangkan bahwa jenis kelamin telah menjadi 0 dan 1, usia, kondisi dari pertanyaan 1 sampai pertanyaan 22. Dari data tersebut maka siap diolah dengan metode K-Means. Setelah data siap diolah akan terbentuk data pada Gambar 5.



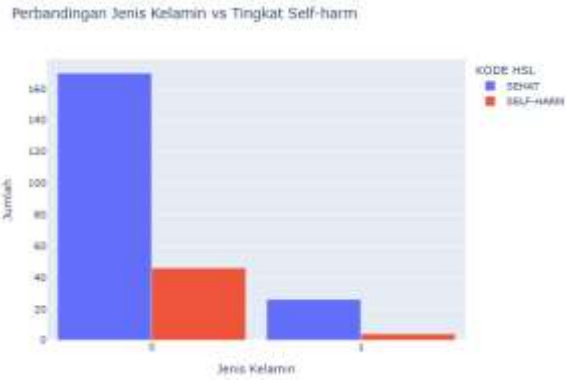
Gambar 5. Sebaran Jenis Kelamin terhadap Tingkat *Self-harm* Mahasiswa Semester Akhir

Gambar 5 menampilkan distribusi antara jenis kelamin (JK) dan tingkat *self-harm* (TS) pada 246 responden. Dari grafik ini terlihat bahwa baik mahasiswa perempuan maupun laki-laki memiliki sebaran pada seluruh tingkat *self-harm* (1–4). Hal ini menunjukkan bahwa tindakan menyakiti diri tidak terbatas pada satu jenis kelamin saja. Terdapat variasi yang cukup jelas pada setiap tingkat, yang mengindikasikan adanya faktor lain di luar jenis kelamin yang turut memengaruhi perilaku self-harm.

Visualisasi ini menjadi dasar awal untuk dilakukan analisis lebih lanjut dengan metode *clustering*. Dengan memetakan distribusi awal ini, dapat melihat bahwa pola persebaran masih acak dan belum menunjukkan pemisahan kelompok yang signifikan. Oleh karena itu, tahap selanjutnya adalah penerapan algoritma *K-Means* untuk menemukan pola laten yang tidak tampak secara langsung.

3.2 K-Means

Tahap pertama analisis clustering dilakukan menggunakan algoritma K-Means dengan jumlah klaster K=2. Tujuan dari tahap ini adalah untuk membedakan mahasiswa yang tergolong sehat dari mereka yang memiliki kecenderungan melakukan self-harm.



Gambar 6. Hasil Clustering Dua Klaster antara Jenis Kelamin dan Tingkat *Self-harm*

Gambar 6 memperlihatkan hasil perbandingan antara jenis kelamin dan hasil pengelompokan dua klaster. Warna biru menunjukkan kelompok "Sehat" dan warna merah menunjukkan kelompok "Self-harm". Terlihat bahwa sebagian besar mahasiswa perempuan (JK=0) berada pada klaster sehat, namun sebagian kecil masuk dalam kelompok self-harm. Mahasiswa laki-laki (JK=1) memiliki distribusi yang serupa, meskipun jumlahnya lebih sedikit.

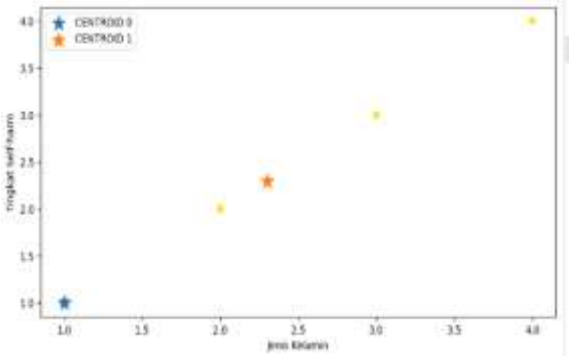
Tabel 1. Crosstab Jenis Kelamin terhadap Hasil Clustering Dua Klaster (K=2)

Kode Hasil JK	Sehat	Self-harm
0	170	46
1	26	4

Dari tabel tersebut terlihat bahwa dari total mahasiswa perempuan **(**JK=0**),** sebanyak 170 orang tergolong dalam klaster "Sehat" dan 46 orang masuk dalam klaster "Self-harm"**.** Sementara itu, dari total mahasiswa laki-laki **(**JK=1**),** terdapat 26 orang dalam klaster "Sehat" dan 4 orang tergolong "Self-harm"**.** Distribusi ini menunjukkan bahwa perempuan cenderung memiliki proporsi yang lebih tinggi dalam kelompok dengan kecenderungan self-harm, meskipun mayoritas dari kedua gender tetap berada pada kondisi sehat. Hasil ini menjadi indikasi awal bahwa faktor gender dapat memiliki pengaruh terhadap kecenderungan perilaku self-harm, sehingga perlu dianalisis lebih lanjut pada tahap klaster

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694

berikutnya.

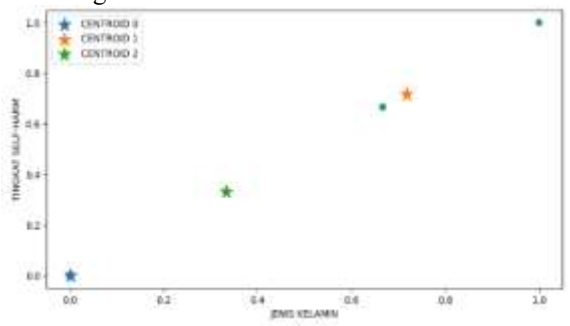


Gambar 7. Visualisasi Dua Klaster Hasil Pengelompokan *K-Means*

Kemudian pada Gambar 7, divisualisasikan hasil *clustering* dua klaster tersebut dengan posisi dua *centroid* yang ditunjukkan oleh tanda bintang berwarna biru dan oranye. Klaster biru menggambarkan kelompok dengan tingkat *self-harm* rendah, sedangkan klaster oranye merepresentasikan kelompok dengan tingkat risiko yang lebih tinggi. Dari visualisasi ini tampak bahwa data telah terbagi ke dalam dua kelompok yang relatif jelas.

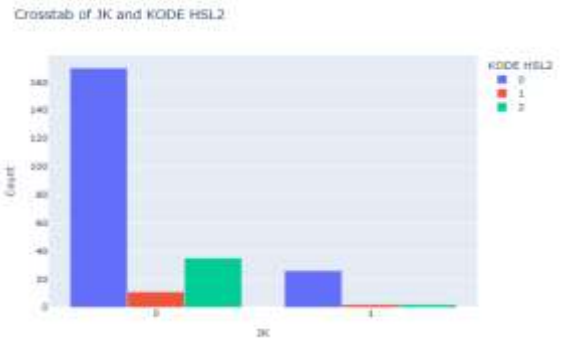
3.3 K-Means Elbow Method

Untuk mendapatkan hasil pengelompokan yang lebih akurat dan representatif, dilakukan penentuan jumlah klaster optimal menggunakan *Elbow Method*. Berdasarkan hasil pengujian, titik siku atau elbow point berada pada nilai K=3. Dengan demikian, jumlah klaster yang paling optimal untuk dataset ini adalah tiga.



Gambar 8. Visualisasi Tiga Klaster Hasil Pengelompokan K-Means

Gambar 8 memperlihatkan hasil pengelompokan dengan tiga klaster. Tiga centroid berwarna biru, oranye, dan hijau menunjukkan kelompok mahasiswa dengan tingkat self-harm berbeda: rendah, sedang, dan tinggi. Pola sebaran yang lebih variatif dibandingkan dua klaster sebelumnya menunjukkan bahwa model K=3 mampu menggambarkan keragaman perilaku mahasiswa secara lebih detail.



Gambar 9. *Crosstab* Jenis Kelamin terhadap Hasil *Clustering* Tiga Klaster

Selanjutnya, hasil crosstab pada Gambar 9 menunjukkan perbandingan antara jenis kelamin dan hasil pengelompokan tiga klaster (KODE HSL2). Terlihat bahwa mayoritas mahasiswa perempuan berada dalam klaster sehat (klaster 0), sedangkan sebagian lainnya tersebar pada klaster 1 dan 2 yang menunjukkan tingkat risiko self-harm sedang hingga tinggi. Pola ini menegaskan bahwa perempuan memiliki tingkat kerentanan psikologis yang sedikit lebih tinggi dibandingkan laki-laki.

Tabel 2. Crosstab Jenis Kelamin terhadap Hasil Clustering Dua Klaster (K=3)

Kode Hasil JK	Sehat	Self-harm ringan	Self-harm berat
0	170	35	11
1	26	2	2

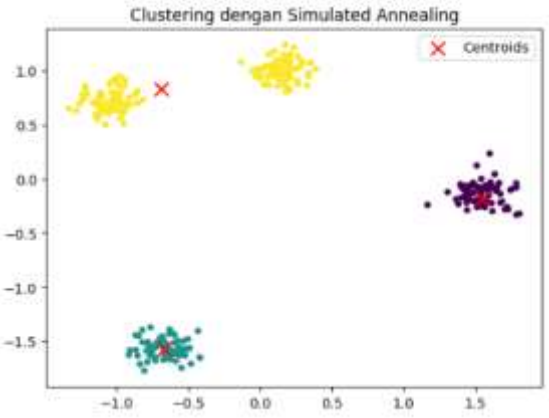
Tabel tersebut menunjukkan bahwa dari total mahasiswa perempuan (JK=0), sebanyak 170 orang berada dalam kategori "Sehat", 35 orang termasuk "*Self-harm* ringan", dan 11 orang tergolong "*Self-harm berat*". Sementara itu, pada kelompok mahasiswa laki-laki (JK=1), terdapat 26 orang dalam kategori "Sehat", 2 orang "*Self-harm* ringan", dan 2 orang "*Self-harm* berat".

Distribusi ini menunjukkan bahwa secara umum, mahasiswa perempuan memiliki kecenderungan lebih tinggi untuk mengalami *self-harm*, baik dalam kategori ringan maupun berat, dibandingkan laki-laki. Meski demikian, sebagian besar responden dari kedua gender masih berada pada kelompok "Sehat".

3.4 Optimasi Menggunakan *Simulated Annealing*

Tahapan berikutnya adalah mengoptimalkan hasil *K-Means* menggunakan algoritma *Simulated Annealing*. Tujuannya adalah untuk menghindari jebakan solusi lokal dan menemukan posisi *centroid* terbaik yang merepresentasikan setiap klaster secara lebih akurat.

INFORMATIKA
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694



Gambar 10. Hasil *Clustering* dengan Optimasi *Simulated Annealing*

Gambar 10 memperlihatkan hasil clustering setelah diterapkan optimasi SA. Tanda X merah menunjukkan posisi centroid baru yang dihasilkan melalui proses penurunan "temperatur" bertahap dalam algoritma. Klaster hasil optimasi terlihat lebih terpisah dan padat, menandakan peningkatan kualitas pengelompokan dibandingkan K-Means murni. Dengan optimasi ini, model berhasil memperjelas batas antar-klaster dan menghasilkan pengelompokan yang lebih reliabel.

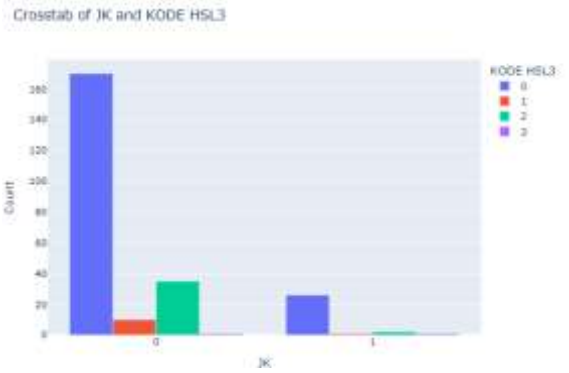
Sebagai tahap lanjutan, jumlah klaster ditingkatkan menjadi K=4 untuk melihat apakah pengelompokan dapat memberikan hasil yang lebih detail.

Gambar 11. Visualisasi Cluster Data dan Centroid

Pada Gambar 11 dapat terlihat kriteria berikut:

d. Klaster 0 (biru): tingkat rendah atau sehat
e. Klaster 1 (ungu): tingkat rendah–menengah
f. Klaster 2 (hijau): tingkat menengah
g. Klaster 3 (merah): tingkat tinggi atau berisiko

Distribusi ini memperlihatkan bahwa proses pengelompokan semakin sensitif dalam membedakan tingkat risiko antar mahasiswa.



Gambar 12. Crosstab Jenis Kelamin terhadap Hasil Clustering Empat Klaster

Sementara itu, Gambar 12 menunjukkan hasil crosstab antara jenis kelamin dan hasil pengelompokan empat klaster (KODE HSL3). Mayoritas mahasiswa perempuan masih berada di klaster sehat (0), namun terdapat sebagian kecil pada klaster risiko menengah hingga tinggi. Mahasiswa laki-laki menunjukkan pola serupa tetapi dengan jumlah responden lebih sedikit.

Tabel 3. Crosstab Jenis Kelamin terhadap Hasil Clustering Dua Klaster (K=4)

Kode Hasil JK	Sehat	Self-harm ringan	Self-harm sedang	Self-harm berat
0	170	10	35	1
1	26	1	2	1

Dari tabel tersebut, terlihat bahwa dari mahasiswa perempuan (JK=0), sebanyak 170 orang berada pada kategori "Sehat", 10 orang termasuk "*Self-harm* ringan", 35 orang "*Self-harm* sedang", dan 1 orang berada pada kategori "*Self-harm* berat". Sementara itu, dari mahasiswa laki-laki (JK=1), terdapat 26 orang dalam kategori "Sehat", 1 orang "*Self-harm* ringan", 2 orang "*Self-harm* sedang", dan 1 orang "*Self-harm* berat".

Hasil ini menunjukkan bahwa meskipun sebagian besar mahasiswa dari kedua jenis kelamin termasuk dalam kelompok Sehat, proporsi mahasiswa perempuan yang masuk dalam kategori "Self-harm sedang" lebih tinggi dibandingkan laki-laki. Hal ini mengindikasikan adanya perbedaan tingkat kerentanan emosional antara gender.

3.5 Hasil Pengujian

Dalam mengukur sejauh mana model algoritma K-Means ini berjalan terhadap data yang digunakan, maka digunakan pengukuran menggunakan *Silhouette Score (SS)*, *Calinski-Harabasz Score (CHS),* dan *Davies-Bouldin Score (DBS)*. Dengan klaster acak, K terbaik dan *Simulated Annealing*.

I N F O R M A T I K A
Jurnal Informatika, Manajemen dan Komputer, Vol. 18 No. 1, Mei 2026
eISSN : 2580-3042
pISSN : 1979-0694

Tabel 4. Perbandingan Pengujian

	Klaster Acak	K Terbaik	*Simulated Annealing*
SS	**0.93**	**0.99**	**1.00**
CHS	**1132.92**	**5753.18**	**1.00**
DBS	**0.34**	**0.19**	**0.00**

Dapat disimpulkan bahwa optimasi *K-Means* dengan menggunakan *Simulated Annealing* berjalan sangat baik terhadap data, ini dibuktikan dengan skor-skor validasi *clustering* setelah menggunakan *Simulated Annealing* (*Silhouette* 1.00, *Calinski-Harabasz* 1.00, *Davies-Bouldin* 0.00) menunjukkan bahwa *clustering* yang dilakukan memiliki kualitas yang sangat baik. Nilai *Silhouette Score* yang mendekati 1 dan *Davies-Bouldin Score* yang mendekati 0 menandakan bahwa hasil klaster sangat terpisah dan memiliki kepadatan internal yang baik. Peningkatan Calinski-Harabasz Score dari 1132,92 menjadi 5753,18 menunjukkan peningkatan rasio dispersi antar-klaster terhadap dalam-klaster yang signifikan. Data-data telah dikelompokkan secara optimal, dengan *cluster-cluster* yang sangat kompak, terpisah jelas satu sama lain, dan membentuk kelompok-kelompok yang sangat terdefinisi.

Secara keseluruhan, hasil pengujian menunjukkan bahwa:

a. Metode K-Means murni sudah mampu memisahkan data ke dalam klaster dasar sehat dan self-harm.
b. Elbow Method (K=3) menambah kedalaman analisis dengan membedakan kategori ringan dan berat.
c. Optimasi dengan Simulated Annealing (K=4) meningkatkan kualitas clustering secara signifikan dengan menghasilkan klaster yang paling representatif terhadap tingkat self-harm mahasiswa.

Dengan demikian, kombinasi K-Means dan Simulated Annealing terbukti lebih unggul dan stabil dalam mengelompokkan mahasiswa berdasarkan kecenderungan perilaku self-harm, dibandingkan metode K-Means standar.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa implementasi metode K-Means Clustering yang dioptimalkan dengan Simulated Annealing efektif dalam mengelompokkan mahasiswa semester akhir berdasarkan pola tindakan self-harm. Optimasi centroid awal menggunakan Simulated Annealing berhasil meningkatkan akurasi dan stabilitas hasil pengelompokan dibandingkan dengan penggunaan K-Means standar. Hal ini ditunjukkan melalui evaluasi menggunakan metrik seperti Silhouette Score, Davies-Bouldin Index, dan Calinski-Harabasz Index, yang menunjukkan performa clustering yang lebih optimal.

Selain itu, penggunaan Simulated Annealing dalam proses pengoptimalan centroid K-Means mampu mengatasi kelemahan K-Means yang sensitif terhadap inisialisasi centroid awal. Dengan demikian, kombinasi metode ini dapat menghasilkan pengelompokan yang lebih efektif dalam memetakan potensi risiko self-harm di kalangan mahasiswa semester akhir.

5. REFERENSI

Asmiyah, S., Madjid, H. I., & Safriyani, R. (2025). Are They Doing Enough for Their Goal? Exploration of Students' Readiness in Writing a Thesis. *IJEE (INDONESIAN JOURNAL OF ENGLISH EDUCATION)*, *12*(1). https://doi.org/10.15408/ijee.v12i1.41525

Hafid, I., & Usop, D. S. (2025). Dinamika Psikologis Perilaku Self-Harm pada Mahasiswa Tingkat Akhir Program Studi Bimbingan Konseling. *Anterior Jurnal*, *24*(2). https://doi.org/https://doi.org/10.33084/anterior.v24i2.9551

Hou, X., Wu, Y., Jiayu Zhao, Luo, J., Jinglong He, Kang, Q., … Luo, J. (2025). Clustering analysis of risk factors for non-suicidal self-injury behaviors in adolescents: a cross-sectional study of western China. *Front Psychiatry*, *22*(16). https://doi.org/10.3389/fpsyt.2025.1436868

Islam, M. R., Banik, S., Rahman, K. N., & Rahman, M. M. (2023). A comparative approach to alleviating the prevalence of diabetes mellitus using machine learning. *Computer Methods and Programs in Biomedicine*, *4*. https://doi.org/https://doi.org/10.1016/j.cmpbup.2023.100113

Oktariani, I. S., Putri, R. M., & Sofah, R. (2021). Tingkat Stress Akademik Mahasiswa dalam Pembelajaran Daring pada Periode Pandemi Covid-19. *Journal of Learning and Instructional Studies*, *1*(1), 17–24. https://doi.org/https://doi.org/10.46637/jlis.v1i1.3

Putra, A. A. (2025). Hybrid Harmony Search Algorithm Dan Simulated Annealing Algorithm Untuk Menyelesaikan Uncapacitated Facility Location Problem. *SCIENCE MATH: Jurnal Ilmu Sains Dan Matematika*, *1*(1), 30–43.

Ridwansyah, R., Riyanto, V., Hamid, A., Rahayu, S., & Purnama, J. J. (2022). Grouping Data in Predicting Infant Mortality Using K-Means and

Decision Tree. *Paradigma*, *24*(2), 168–174. https://doi.org/10.31294/paradigma.v24i2.1399

Shaikh, M. S., Dong, X., Zheng, G., Wang, C., & Lin, Y. (2024). An Improved Expeditious Meta-Heuristic Clustering Method for Classifying Student Psychological Issues with Homogeneous Characteristics. *Mathematics*, *12*(11). https://doi.org/https://doi.org/10.3390/math12111620

Suhardjono, Sugiarto, H., Yuliandari, D., Sudradjat, A., & Rohimah, L. (2023). Classifying Half-Unemployment Levels in Indonesian Provinces: A K-Means Approach for Informed Policy Decisions. *Penelitian Ilmu Komputer Sistem Embedded and Logic*, *11*(2). https://doi.org/https://doi.org/10.33558/piksel.v11i2.7390

Tarigan, T., & Apsari, N. C. (2024). Perilaku Self-Harm atau Melukai Diri Sendiri yang Dilakukan oleh Remaja (Self-Harm or Self-Injuring Behavior by Adolescents). *Jurnal Pekerjaan Sosial*, *4*(2). https://doi.org/https://doi.org/10.24198/focus.v4i2.31405

Yang, H., Wen, X., & Geng, P. (2024). An Optimisation Strategy for Electric Vehicle Charging Station Layout Incorporating Mini Batch K-Means and Simulated Annealing Algorithms. *Journal on Aritficial Intelligence*, *6*. https://doi.org/10.32604/jai.2024.056303

Yang, J., Wang, Y.-K., Yao, X., & Lin, C.-T. (2021). Adaptive Initialization Method for K-Means Algorithm. *Machine Learning and Artificial Intelligence*, *4*. https://doi.org/https://doi.org/10.3389/frai.2021.740817

Yang, R., Danlin Li, Run Tian, Jie Hu, Yanni Xue, Xuexue Huang, … Zang, S. (2021). Association of Unhealthy Behaviors with Self-Harm in Chinese Adolescents: A Study Using Latent Class Analysis. *Trauma Care*, *1*(2), 75–86. https://doi.org/https://doi.org/10.3390/traumacare1020008

Yennimar, Leonardi, W., Weide, H., Cantona, D., & Gani Mores Hutagalung. (2024). Comparison of data mining algorithms (random forest, C4.5, catboost) based on adaptive boosting in predicting diabetes mellitus. *Jurnal Teknik Informatika C.I.T Medicom*, *16*(3), 01–12. Retrieved from www.medikom.iocspublisher.orgjournalhomepage:www.medikom.iocspublisher.org