

Pengelompokan Tingkat Kesejahteraan dan Tekanan Psikologis Mahasiswa Menggunakan Mental Health Inventory-38 dengan Algoritma K-Means

Yadullah Asy-syakiri¹, Elvia Budianita^{2,*}, Iwan Iskandar³, Fadhilah Syafria⁴, Yuli Widiningsih⁵

¹ Fakultas Sains dan Teknologi, Prodi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

² Fakultas Psikologi, Program Studi Psikologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹12250115109@students.uin-suska.ac.id, ²elvia.budianita@uin-suska.ac.id, ³iwan.iskandar@uin-suska.ac.id,

⁴fadhilah.syafria@uin-suska.ac.id, ⁵yuli.widiningsih@uin-suska.ac.id

Email Penulis Korespondensi: elvia.budianita@uin-suska.ac.id

Abstrak—Ketiadaan pemetaan objektif kondisi psikologis mahasiswa STEM menjadi hambatan dalam merancang intervensi kesehatan mental yang tepat sasaran. Penelitian ini bertujuan mengidentifikasi pola kesehatan mental mahasiswa Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau angkatan 2022–2025 menggunakan algoritma K-Means Clustering berbasis instrumen Mental Health Inventory-38 (MHI-38) yang diadaptasi ke Bahasa Indonesia dan divalidasi psikolog ahli. Dari 559 data yang dikumpulkan, 522 data valid digunakan setelah data cleaning. Clustering diterapkan dengan dua metode jarak Euclidean dan Manhattan pada $k=2$ hingga 10, dievaluasi menggunakan Davies-Bouldin Index (DBI) dan Silhouette Coefficient. Konfigurasi optimal diperoleh pada $k=2$: Euclidean menghasilkan DBI 2,03 dan Silhouette 0,18, sedangkan Manhattan menghasilkan DBI 2,19 dan Silhouette 0,20. Nilai Silhouette yang rendah (0,18–0,20) menunjukkan bahwa struktur klaster masih lemah. Kondisi ini kemungkinan dipengaruhi oleh heterogenitas program studi dan distribusi responden yang tidak merata sehingga hasil perlu diinterpretasikan secara hati-hati. Meski demikian, pengelompokan secara konsisten membentuk dua klaster yang bermakna secara psikologis: Klaster 0 terdiri dari 334 mahasiswa (64%) berkategori kesejahteraan psikologis (skor rata-rata 3,82), dan Klaster 1 terdiri dari 188 mahasiswa (36%) berkategori tekanan psikologis (skor rata-rata 2,98). Teknik Informatika menjadi satu-satunya prodi yang didominasi Klaster 1 (61,8%), sedangkan angkatan 2022 tercatat paling stabil secara psikologis (72,4% di Klaster 0). Temuan ini berkontribusi sebagai peta distribusi kesehatan mental berbasis data empiris yang dapat menjadi landasan bagi dekanat dan unit konseling dalam merancang program intervensi psikologis yang bertarget serta kebijakan akademik yang berpihak pada kesejahteraan mahasiswa.

Kata Kunci: K-Means Clustering; Kesehatan Mental Mahasiswa; Mental Health Inventory-38; Davies-Bouldin Index; Silhouette Coefficient

Abstract—The lack of objective mapping of the psychological conditions of STEM students poses a barrier to designing targeted mental health interventions. This study aims to identify mental health patterns among students in the Faculty of Science and Technology at UIN Sultan Syarif Kasim Riau, class of 2022–2025, using the K-Means Clustering algorithm based on the Mental Health Inventory-38 (MHI-38), which was adapted into Indonesian and validated by expert psychologists. Of the 559 data points collected, 522 valid data points were used after data cleaning. Clustering was applied using two distance measures Euclidean and Manhattan at $k=2$ through 10, and evaluated using the Davies-Bouldin Index (DBI) and the Silhouette Coefficient. The optimal configuration was obtained at $k=2$: the Euclidean method yielded a DBI of 2.03 and a Silhouette Coefficient of 0.18, while the Manhattan method yielded a DBI of 2.19 and a Silhouette Coefficient of 0.20. The low Silhouette Coefficient values (0.18–0.20) indicate that the cluster structure is still weak. This condition is likely influenced by the heterogeneity of the study programs and the uneven distribution of respondents, so the results must be interpreted with caution. Nevertheless, the clustering consistently formed two psychologically meaningful clusters: Cluster 0 consists of 334 students (64%) categorized as having psychological well-being (average score of 3.82), and Cluster 1 consists of 188 students (36%) categorized as experiencing psychological stress (average score of 2.98). Computer Science was the only program dominated by Cluster 1 (61.8%), while the class of 2022 was found to be the most psychologically stable (72.4% in Cluster 0). These findings contribute to an empirically based map of mental health distribution that can serve as a foundation for the dean's office and counseling units in designing targeted psychological intervention programs and academic policies that prioritize student well-being.

Keywords: K-Means Clustering; Student Mental Health; Mental Health Inventory-38; Davies-Bouldin Index; Silhouette Coefficient

1. PENDAHULUAN

Kesehatan mental mahasiswa telah menjadi isu global yang kian mendesak, khususnya pasca-pandemi yang memperparah tekanan psikologis pada kelompok usia produktif. Kenaikan prevalensi gangguan mental secara global sebesar 25%, dengan depresi dan kecemasan sebagai kondisi yang paling dominan [1]. Lebih dari delapan juta mahasiswa di berbagai negara dilaporkan mengalami gejala depresi ringan (35%), kecemasan ringan (40%), stres psikologis (36%), serta gangguan tidur klinis (41%) [2]. Di Indonesia, kondisi ini tercermin secara empiris: sebanyak 157 mahasiswa (66%) dari 238 responden mengalami kecemasan kategori sedang, 81 mahasiswa (75%) dari 108 mengalami depresi ringan, dan 44 mahasiswa (81%) dari 54 berada pada kategori stres ringan [3]. Kondisi ini dirasakan lebih intens oleh mahasiswa bidang Sains, Teknologi, Rekayasa, dan Matematika (STEM) yang menghadapi tuntutan kurikulum berat dan iklim akademik yang kompetitif [4]. Karakteristik kurikulum STEM yang mencakup jadwal praktikum padat, proyek berbasis penalaran logika tingkat tinggi, serta ekspektasi kompetensi teknis yang ketat menjadikan mahasiswa pada rumpun ini lebih rentan terhadap akumulasi tekanan psikologis dibandingkan mahasiswa di bidang lain [5]. Apabila tidak ditangani secara tepat, tekanan tersebut berpotensi memicu *burnout*, penurunan prestasi, hingga keterlambatan kelulusan.

Kesehatan mental di 1 pola tersembunyi dari dataset berskala besar yang tidak dapat diidentifikasi melalui analisis statistik konvensional. K-Means dipilih karena memiliki efisiensi komputasi yang tinggi, skalabilitas yang baik terhadap



data berdimensi banyak, serta kemampuannya dalam mengidentifikasi struktur pengelompokan secara andal. Meskipun metode psikometrika seperti Latent Class Analysis (LCA) secara teori lebih sesuai untuk data ordinal berdimensi terstruktur, penelitian ini memilih K-Means dengan pertimbangan tujuan penelitian bersifat eksploratif-deskriptif untuk mendukung pengambilan keputusan kebijakan institusional, bukan konfirmasi struktur laten, K-Means telah terbukti secara empiris menghasilkan pengelompokan yang stabil dan interpretatif pada data survei psikologis serupa, dan perbandingan dua metrik evaluasi (DBI dan Silhouette) serta dua metode jarak (Euclidean dan Manhattan) dilakukan untuk mengkompensasi keterbatasan inherent K-Means dalam menangani data ordinal. Pendekatan ini diterapkan untuk memetakan pola kesehatan mental mahasiswa Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau secara objektif dan *data-driven*. Instrumen yang digunakan adalah *Mental Health Inventory-38* (MHI-38) yang dikembangkan oleh [7], mencakup dua dimensi utama yaitu tekanan psikologis (*psychological distress*) dan kesejahteraan psikologis (*psychological well-being*).

Sejumlah penelitian terdahulu telah mengeksplorasi pendekatan berbasis data mining untuk analisis kesehatan mental mahasiswa, namun masing-masing menyisakan celah yang signifikan. Sebuah kerangka kerja konseptual penerapan data mining dalam pendidikan kesehatan mental mahasiswa telah didemonstrasikan, namun tidak mengimplementasikan klusterisasi dengan instrumen tervalidasi, sehingga tidak menghasilkan profil kluster yang dapat digunakan secara operasional [8]. Upaya klasifikasi kondisi kesehatan mental mahasiswa ke dalam tiga kategori menggunakan K-Means telah dilakukan, namun tanpa instrumen baku yang tervalidasi secara psikometrik dan tanpa analisis silang berbasis data demografis, sehingga hasilnya sulit digeneralisasi dan tidak memberikan informasi spesifik tentang kelompok mahasiswa yang paling rentan [9]. Pengelompokan pola kesehatan mental mahasiswa menggunakan K-Means dengan evaluasi Silhouette Coefficient telah berhasil dilakukan, namun penelitian tersebut masih berfokus pada pemetaan pola umum dan belum menggunakan instrumen psikometrik yang diadaptasi secara budaya lokal [10]. Penerapan K-Means dalam konteks kesehatan mental mahasiswa pendidikan tinggi dengan pendekatan *fairness-aware* juga telah dikaji, namun studi tersebut lebih menekankan aspek keadilan pengelompokan daripada pemetaan kondisi psikologis berbasis instrumen survei yang tervalidasi [11]. Secara kolektif, keempat penelitian ini belum ada yang menggabungkan instrumen MHI-38 yang diadaptasi secara budaya, populasi mahasiswa STEM di perguruan tinggi Islam Indonesia, dan analisis demografis lintas angkatan serta program studi secara bersamaan. Meskipun sejumlah penelitian terdahulu telah menerapkan K-Means untuk mengelompokkan data kesehatan mental mahasiswa, belum ada studi yang secara spesifik memetakan profil kerentanan psikologis mahasiswa pada perguruan tinggi berbasis Islam (UIN) di Indonesia menggunakan MHI-38 yang telah diadaptasi secara kultural, serta membandingkan pola antar-angkatan dan antar-program studi secara simultan dalam satu institusi. Selain itu, studi terdahulu umumnya tidak memvalidasi kualitas clustering secara komparatif menggunakan lebih dari satu metrik evaluasi maupun lebih dari satu metode jarak.

Berdasarkan penelitian terkait sebelumnya, dapat disimpulkan bahwa algoritma K-Means efektif dalam mengklasifikasikan data kesehatan mental mahasiswa. Kendati demikian, riset yang menggunakan algoritma K-Means pada mahasiswa Universitas Islam Negeri Sultan Syarif Kasim Riau dengan instrumen kuesioner MHI-38 yang diadaptasi ke dalam Bahasa Indonesia dan dimodifikasi secara spesifik masih minim. Penggabungan populasi lokal dengan instrumen yang telah melalui proses adaptasi budaya menawarkan pendekatan baru dalam analisis pengelompokan pola kesehatan mental, yang pada gilirannya memperkaya relevansi dan kesesuaian kontekstual temuan penelitian. Data yang dikumpulkan berasal dari mahasiswa Fakultas Sains dan Teknologi, menggunakan kuesioner MHI-38 yang telah diadaptasi ke Bahasa Indonesia dan dimodifikasi oleh seorang psikolog ahli, yaitu Yuli Widiningsih, S.Psi., M.Psi.

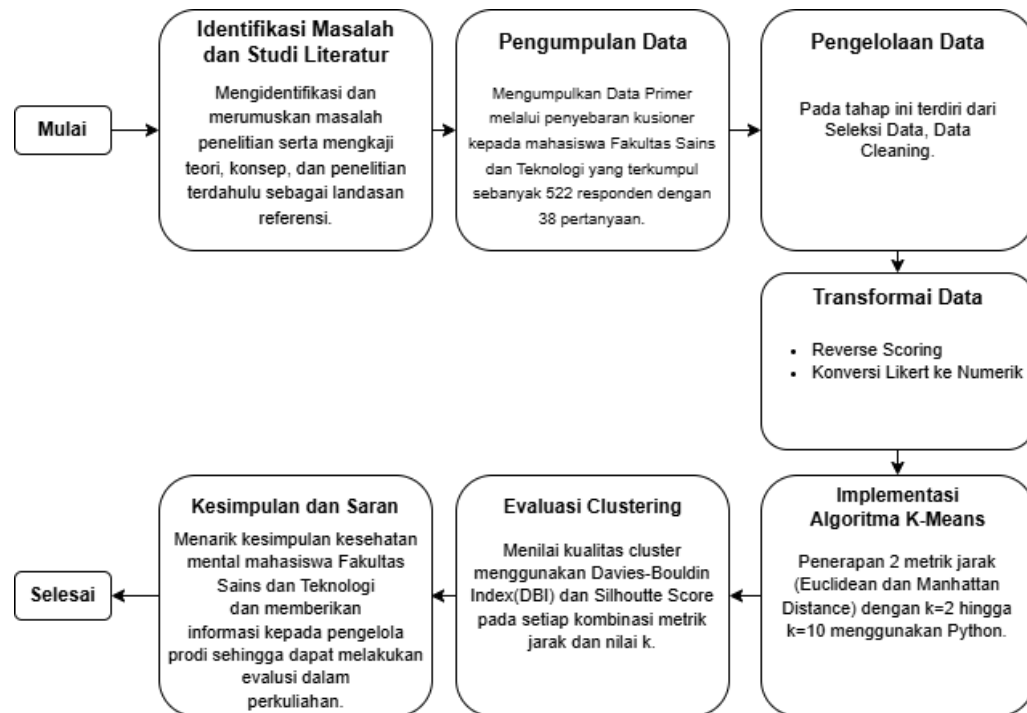
Penelitian ini memberikan kontribusi berupa pemetaan awal kondisi kesehatan mental mahasiswa Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau menggunakan algoritma K-Means Clustering berdasarkan instrumen Mental Health Inventory-38 (MHI-38), sehingga dapat menjadi acuan strategis bagi pihak dekanat dan unit konseling universitas dalam merancang program intervensi psikologis dini, menetapkan prioritas layanan konseling, serta merumuskan kebijakan akademik yang lebih berpihak pada kesejahteraan mental mahasiswa.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian ini mengikuti kerangka kerja Knowledge Discovery in Databases (KDD) yang terdiri atas tahapan identifikasi masalah, studi literatur, pengumpulan data, seleksi data, preprocessing, transformasi data, implementasi K-Means Clustering, dan evaluasi. Alur keseluruhan proses penelitian disajikan secara sistematis pada Gambar 1 berikut.

Gambar 1 mengilustrasikan delapan tahapan penelitian yang berjalan secara sekuensial, dimulai dari identifikasi masalah hingga evaluasi hasil kluster. Setiap tahapan saling bergantung sehingga output satu tahap menjadi input bagi tahap berikutnya, memastikan proses analisis berjalan secara sistematis dan dapat direplikasi.



Gambar 1. Metodologi Penelitian Pengelompokan Kesejahteraan dan Tekanan Psikologis Mahasiswa

2.2 Identifikasi Masalah

Permasalahan utama penelitian ini adalah belum tersedianya pemetaan kondisi psikologis mahasiswa Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau yang bersifat data-driven dan resolusi tinggi. Data yang tersedia saat ini hanya menyajikan angka per indikator tanpa mengungkap pola kerentanan laten yang tersembunyi di balik distribusi data, sehingga intervensi yang dirancang cenderung reaktif dan tidak tertarget.

2.3 Studi Literatur

Studi literatur dilakukan untuk memperoleh pemahaman komprehensif terhadap konsep kesehatan mental mahasiswa, instrumen pengukuran MHI-38, metode K-Means Clustering, serta teknik evaluasi kluster yang relevan. Selama proses analisis data, studi ini mengadopsi kerangka kerja KDD yang meliputi fase-fase pemilihan data, rekayasa, transformasi, permodelan, serta penilaian, sebagaimana telah diterapkan dalam berbagai penelitian analisis kesehatan mental mahasiswa menggunakan teknologi clustering berbasis data mining.

2.4 Pengumpulan Data

Data dikumpulkan melalui penyebaran kuesioner MHI-38 yang diadaptasi ke Bahasa Indonesia dan telah divalidasi oleh psikolog ahli (Yuli Widiningsih, S.Psi., M.Psi.) kepada mahasiswa Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau angkatan 2022–2025 melalui Google Form. Instrumen MHI-38 mencakup 38 butir pertanyaan yang mengukur dua dimensi utama: tekanan psikologis (psychological distress, item 1–23) dan kesejahteraan psikologis (psychological well-being, item 24–38), dengan skala respons Likert lima poin (Sangat Sering hingga Tidak Pernah).

Tekanan psikologis merupakan kondisi individu yang menunjukkan keadaan kesehatan mental yang kurang optimal. Mendefinisikan psychological distress sebagai kondisi penderitaan emosional yang ditandai dengan munculnya gejala-gejala seperti depresi dan kecemasan yang dirasakan oleh individu akibat ketidakmampuan dalam menghadapi tekanan kehidupan sehari-hari[12]. Secara umum, tekanan psikologis terdiri dari tiga aspek utama yaitu *Anxiety*, *Depression*, dan *Lost of behavioral/emotional control*. Sedangkan kesejahteraan psikologis merupakan kondisi individu yang ditandai dengan kemampuan untuk merealisasikan potensi diri, mengelola tekanan kehidupan sehari-hari secara wajar, bekerja secara produktif, serta memberikan kontribusi positif kepada lingkungan sosial atau komunitasnya. Psychological well-being adalah kondisi mental di mana individu mampu mencapai potensi penuhnya, produktif dan inovatif dalam pekerjaannya, serta mampu mengelola berbagai stresor kehidupan sehari-hari secara efektif[13]. Secara umum, kesejahteraan psikologis terdiri dari dua aspek utama yaitu *Emotional Tie* dan *General positive effect*.

2.5 Seleksi Data

Pada tahap seleksi data, atribut awal yang dikumpulkan meliputi nama, jenis kelamin, Nomor Induk Mahasiswa (NIM), angkatan, program studi, dan respons terhadap pertanyaan dalam kuesioner MHI-38 yang telah diselesaikan oleh partisipan. Kajian ini mengadopsi skala Likert untuk pengukuran, yang menawarkan lima opsi respons: Sangat Sering, Sering, Kadang-kadang, Jarang, dan Tidak Pernah. Selama proses ini, data juga diklasifikasikan ke dalam dua kategori: UF (Unfavorable), yang mewakili respons dengan kecenderungan negatif, dan F (Favorable) yang mencerminkan respons

dengan kecenderungan positif. Kuesioner MHI-38, yang berisi 38 pertanyaan, disajikan pada Tabel 1. Setiap pertanyaan dirancang berdasarkan indikator dari kuesioner MHI-38 yang telah mengalami modifikasi.

Tabel 1. Instrumen MHI-38

No	Pertanyaan	SS	S	K	J	T
				K		P
1	Seberapa sering kamu merasa sangat gugup? (UF)					
2	Seberapa sering kamu terganggu oleh rasa gugup? (UF)					
3	Seberapa sering kamu merasa tegang atau sangat tertekan? (UF)					
4	Seberapa sering kamu merasa cemas atau khawatir? (UF)					
5	Seberapa sering kamu kesulitan untuk menenangkan diri? (UF)					
6	Seberapa sering kamu merasa gugup hingga gemetar? (UF)					
7	Seberapa sering kamu merasa gelisah? (UF)					
8	Seberapa sering kamu merasa resah, kesal, atau bingung? (UF)					
9	Seberapa sering kamu merasa tangan gemetar saat melakukan aktivitas? (UF)					
10	Seberapa sering kamu merasa suasana hati murung atau sering memikirkan hal-hal negatif? (UF)					
11	Seberapa sering kamu merasa semangat hidup rendah? (UF)					
12	Seberapa sering kamu merasa terpuruk dan sedih? (UF)					
13	Seberapa sering kamu merasa tertekan? (UF)					
14	Seberapa sering kamu merasa mengalami tekanan, stres, atau beban berat? (UF)					
15	Seberapa sering kamu merasa hilang kendali terhadap perilaku, pikiran, dan perasaan? (UF)					
16	Seberapa sering datang kekhawatiran mengenai hilangnya kendali diri? (UF)					
17	Seberapa sering kamu merasa emosi tidak stabil? (UF)					
18	Seberapa sering kamu merasa tidak ada yang terasa menyenangkan? (UF)					
19	Seberapa sering kamu merasa ingin menangis? (UF)					
20	Seberapa sering kamu merasa lebih baik jika meninggal? (UF)					
21	Seberapa sering kamu merasa sangat terpuruk? (UF)					
22	Seberapa sering kamu memikirkan untuk mengakhiri hidup? (UF)					
23	Seberapa sering kamu tidak menantikan masa depan? (UF)					
24	Seberapa sering kamu merasa bahagia? (F)					
25	Seberapa sering kamu merasa puas, senang, atau gembira? (F)					
26	Seberapa sering kamu merasa kehidupan sehari-hari terasa menarik? (F)					
27	Seberapa sering kamu merasa tenang dan damai? (F)					
28	Seberapa sering kamu merasa ceria dan bersemangat? (F)					
29	Seberapa sering kamu merasa benar-benar menikmati berbagai aktivitas? (F)					
30	Seberapa sering kamu menyadari dan menghargai hal-hal yang menyenangkan? (F)					
31	Seberapa sering kamu menjalani hidup sebagai sebuah petualangan yang bermakna? (F)					
32	Seberapa sering kamu mengharapkan hal-hal menarik di masa depan? (F)					
33	Seberapa sering kamu bangun tidur dengan perasaan segar dan beristirahat cukup? (F)					
34	Seberapa sering masa depan terasa penuh harapan dan menjanjikan? (F)					
35	Seberapa sering kamu merasa dicintai dan diinginkan? (F)					
36	Seberapa sering kamu merasa memiliki hubungan yang penuh kasih? (F)					
37	Seberapa sering kamu merasa utuh secara emosional? (F)					
38	Seberapa sering kamu merasa dicintai sepenuh hati? (F)					

2.6 Data Preprocessing

Tahapan pra-pemrosesan data meliputi pembersihan data agar sesuai untuk analisis lanjutan. Setiap tahapan ini dimaksudkan untuk memberikan wawasan yang lebih komprehensif mengenai kesehatan mental para mahasiswa di Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.

2.6.1 Data Cleaning

Data cleaning dilakukan untuk memastikan konsistensi dan kelengkapan dataset sebelum analisis. Proses ini mencakup pembersihan data data duplikat, normalisasi menggunakan metode Min-Max Scaling, perbaikan inkonsistensi format jawaban, serta eliminasi entri dengan respons tidak valid [14]. Tahap ini menghasilkan pengurangan dari 559 data menjadi 522 data yang bersih dan siap dianalisis. Dalam studi ini, pembersihan data dilaksanakan dengan mengeliminasi data



responden yang tidak memenuhi spesifikasi yang ditentukan, yaitu mahasiswa Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Kasim Riau angkatan 2022–2025.

2.7 Transformasi Data

Tahapan ini dilaksanakan dengan mengonversi informasi menjadi representasi numerik, sehingga seluruh data menjadi kuantitatif dan konsisten dalam formatnya. Konversi ini dirancang untuk menyederhanakan perhitungan jarak antar data oleh algoritma dan memfasilitasi proses pengelompokan data. Setiap item pertanyaan dalam kuesioner diklasifikasikan ke dalam kategori Mendukung (F) dan Tidak Mendukung (UF). Untuk pertanyaan Mendukung, penilaian skor berkisar dari 1 untuk respons "Tidak Pernah" hingga 5 untuk respons "Sangat Sering". Sebaliknya, untuk pertanyaan Tidak Mendukung, skor diberikan secara terbalik, dimulai dari 1 untuk "Sangat Sering" hingga 5 untuk "Tidak Pernah".

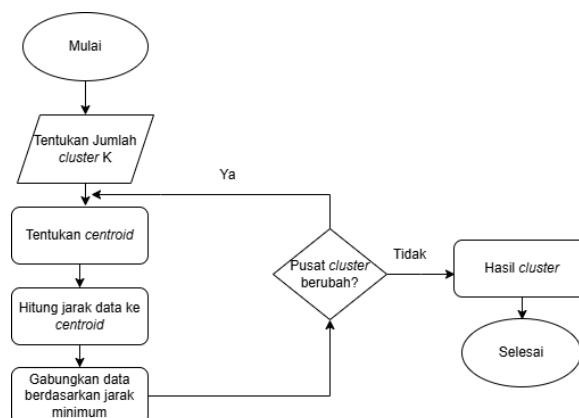
Mengingat algoritma K-Means berbasis perhitungan jarak (distance-based), perbedaan skala antar atribut berpotensi mendistorsi posisi centroid dan menurunkan kualitas kluster [15]. Meskipun seluruh atribut dalam penelitian ini telah berada pada rentang yang seragam (1–5) sebagai hasil transformasi Likert, normalisasi Min-Max diterapkan untuk memastikan seluruh atribut berkontribusi secara proporsional dalam perhitungan jarak Euclidean dan Manhattan. Hasil pengubahan skor ini dapat dilihat pada Tabel 2 berikut.

Tabel 2. Data Transformation

No	Jawaban	UF	F
1	Sangat Sering	1	5
2	Sering	2	4
3	Kadang-kadang	3	3
4	Jarang	4	2
5	Tidak Pernah	5	1

2.8 Implementasi K-Means Clustering

K-Means merupakan salah satu metode *unsupervised learning* yang digunakan untuk mengolah data tanpa label atau klasifikasi awal. Metode ini bekerja dengan mengenali pola-pola kemiripan yang terdapat dalam data, kemudian mengelompokkan data tersebut ke dalam cluster tertentu berdasarkan karakteristik yang serupa [16]. Namun, K-Means memiliki keterbatasan, antara lain penentuan jumlah kluster (k) yang tidak mudah, pemilihan pusat kluster awal, dan sensitivitas terhadap outlier [17]. Tahapan penerapan metode K-Means pada penelitian ini disajikan secara sistematis melalui flowchart pada Gambar 2 berikut.



Gambar 2. Flowchart Metode K-Means

Implementasi algoritma K-Means Clustering dilakukan melalui serangkaian tahapan yang berjalan secara iteratif. Proses dimulai dengan menentukan jumlah kluster (k) berdasarkan kebutuhan analisis atau hasil metode evaluasi, dilanjutkan dengan penetapan posisi centroid awal secara acak sebagai titik pusat representasi setiap kluster. Selanjutnya, algoritma menghitung jarak antara setiap titik data terhadap masing-masing centroid, kemudian mengalokasikan setiap data ke kluster dengan nilai jarak minimum (nearest centroid). Setelah seluruh data berhasil dikelompokkan, posisi centroid baru pada masing-masing kluster dihitung ulang menggunakan persamaan berikut:

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} x_q \quad (1)$$

Keterangan: μ_k merupakan centroid pada cluster ke- (k) , N_k adalah jumlah data pada cluster ke- (k) , dan x_q merupakan data ke- (q) yang berada dalam *cluster*. Apabila posisi centroid baru berbeda dari posisi sebelumnya, algoritma kembali menghitung jarak seluruh data terhadap centroid terbaru dan melakukan pengelompokan ulang. Proses iterasi ini terus berulang hingga posisi centroid tidak lagi mengalami perubahan yang signifikan, yang menandakan konvergensi telah tercapai. Pada titik tersebut, algoritma berhenti dan menghasilkan pembagian akhir seluruh data ke dalam sejumlah k kluster yang telah terbentuk secara optimal.

Dalam penelitian ini metode perhitungan jarak yang digunakan menggunakan Euclidean dan Manhattan Distance. Euclidean Distance dalam ranah analisis data, salah satu metode paling umum untuk mengukur jarak adalah jarak Euclidean. Pendekatan ini diaplikasikan untuk mengidentifikasi jarak yang memisahkan dua titik data, seperti x dan y , dalam ruang berdimensi d , yang proses perhitungannya dapat dirujuk pada persamaan 2.

$$d = \sqrt{\{(x_2 - x_1)^2 + (y_2 - y_1)^2\}} \quad (2)$$

Keterangan: d merupakan jarak Euclidean, x adalah data, dan y merupakan centroidnya. Perhitungan jarak kedua menggunakan Manhattan, jarak dikalkulasikan dengan menggunakan dua titik koordinat dari entitas yang sedang dikomparasi. Berikut ini adalah formulasi yang digunakan dalam perhitungan jarak Manhattan untuk menguji banding kedua metodologi pengukuran jarak yang diterapkan:

$$d = \sum_{i=1}^n |x_i - y_i| \quad (3)$$

Keterangan: d merupakan jarak Manhattan antara dua titik/objek, n adalah jumlah atribut/variable/dimensi yang dibandingkan, dan $|x_i - y_i|$ merupakan selisih absolut antara atribut dari dua objek.

3. HASIL DAN PEMBAHASAN

3.1 Analisis Kebutuhan

Dalam penelitian ini, mahasiswa dikelompokkan berdasarkan kondisi kesehatan mental yang mereka alami. Pengumpulan data dilakukan menggunakan instrumen *Mental Health Inventory-38* yang dikemukakan oleh [7]. Kemudian divalidasi ahli oleh Yuli Widiningsih, S.Psi., M.Psi. Instrumen tersebut digunakan untuk memperoleh data kesehatan mental mahasiswa Fakultas Sains dan Teknologi angkatan 2022–2025 melalui penyebaran kuesioner. Dataset penelitian terdiri dari 522 responden dengan 38 butir pertanyaan yang mengukur dua dimensi utama kesehatan mental, yaitu kesejahteraan psikologis (*psychological well-being*) dan tekanan psikologis (*psychological distress*). Hasil pada tahap ini menjadi landasan penting dalam memahami kondisi psikologis dan kebutuhan akademik mahasiswa secara lebih menyeluruh, sehingga dapat digunakan sebagai dasar dalam penyusunan program dukungan yang lebih efektif.

3.2 Seleksi Data

Sumber data pada penelitian ini berasal dari kuesioner yang diisi oleh mahasiswa Fakultas Sains dan Teknologi angkatan 2022–2025. Kuesioner tersebut terdiri dari 38 pertanyaan yang digunakan untuk memperoleh informasi terkait kondisi kesehatan mental mahasiswa. Data yang telah terkumpul selanjutnya melalui tahap seleksi guna memastikan bahwa hanya data yang lengkap dan valid yang digunakan pada proses analisis berikutnya. Hasil dari proses seleksi data tersebut disajikan pada Tabel 3 sebagai dataset yang siap digunakan dalam proses pengelompokan menggunakan algoritma clustering.

Tabel 3. Data Selection

No	Q1 (UF)	Q2 (UF)	Q3 (UF)	Q4 (UF)	Q5 (UF)	Q6 (UF)	Q7 (UF)	Q8 (UF)	Q9 (UF)	...	Q38 (F)
1	Sering	Sering	Jarang	Sering	Jarang	Tidak Pernah	Sering	Sering	Tidak Pernah	...	Jarang
2	Sering	Jarang	Jarang	Sering	Jarang	Sering	Jarang	Jarang	Jarang	...	Kadang-kadang
3	Sering	Jarang	Jarang	Sering	Sering	Jarang	Jarang	Jarang	Jarang	...	Sangat Sering
4	Sangat Sering	Sangat Sering	Sangat Sering	Sangat Sering	Sangat Sering	Sangat Sering	Sangat Sering	Sangat Sering	Sangat Sering	...	Sangat Sering
5	Sering	Sering	Sering	Sangat Sering	Sangat Sering	Sering	Sering	Sering	Sangat Sering	...	Kadang-kadang
...
518	Kadang- Kadang	Tidak Pernah	Jarang	Kadang- Kadang	Tidak Pernah	Jarang	Tidak Pernah	Jarang	Jarang	...	Jarang
519	Jarang	Jarang	Tidak Pernah	Tidak Pernah	Tidak Pernah	Tidak Pernah	Jarang	Tidak Pernah	Tidak Pernah	...	Kadang-kadang
520	Jarang	Kadang- Kadang	Tidak Pernah	Kadang- Kadang	Jarang	Jarang	Kadang- Kadang	Jarang	Tidak Pernah	...	Sering
521	Jarang	Jarang	Tidak Pernah	Tidak Pernah	Jarang	Jarang	Jarang	Tidak Pernah	Tidak Pernah	...	Kadang-kadang
522	Jarang	Jarang	Tidak Pernah	Tidak Pernah	Tidak Pernah	Tidak Pernah	Tidak Pernah	Tidak Pernah	Tidak Pernah	...	Kadang-kadang

3.3 Data Preprocessing

Tahap ini meliputi proses persiapan, pembersihan, dan transformasi data mentah agar data siap untuk dianalisis lebih lanjut. Tahapan tersebut penting dilakukan untuk memastikan data memiliki tingkat akurasi dan konsistensi yang baik sebelum memasuki proses pemodelan.



3.3.1 Data Cleaning

Tahap ini dilakukan untuk membersihkan data dengan cara menghapus data duplikat, memperbaiki kesalahan pengisian, serta memastikan data yang digunakan memiliki konsistensi dan kelengkapan yang baik sebelum proses analisis dilakukan. Proses ini bertujuan untuk mempermudah pengolahan data serta meningkatkan tingkat akurasi hasil analisis clustering pada penelitian. Dalam penelitian ini, jumlah data yang semula sebanyak 559 data berkurang menjadi 522 data setelah proses pembersihan dilakukan. Dengan pelaksanaan data cleaning secara menyeluruh, kualitas dataset menjadi lebih baik sehingga hasil analisis clustering yang diperoleh dapat lebih representatif, relevan, dan dapat dipercaya dalam mendukung kesimpulan penelitian.

3.4 Transformasi Data

Pada tahap ini, dilakukan penyesuaian format data agar sesuai dengan kebutuhan dan standar penelitian. Setelah data dibersihkan dari jawaban yang tidak relevan, proses berikutnya adalah transformasi data. Tahap ini dilakukan dengan mengubah setiap jawaban pada kuesioner ke dalam bentuk angka (numerik) berdasarkan matriks skala penilaian yang telah ditentukan pada tabel berikut, mengingat proses klasterisasi mensyaratkan seluruh atribut data berada dalam format numerik sebelum dapat digunakan dalam analisis lebih lanjut[18].

Nilai numerik tersebut merepresentasikan tingkat kecenderungan jawaban responden terhadap setiap pertanyaan, mulai dari kategori sangat sering, sering, kadang-kadang, jarang, hingga tidak pernah. Transformasi data menjadi tahapan penting karena bertujuan untuk memastikan data yang akan dianalisis memiliki validitas yang baik dan sesuai dengan kebutuhan penelitian. Adapun hasil transformasi data pada penelitian ini disajikan secara rinci pada Tabel 4 berikut.

Tabel 4. Data Transformation Kesehatan Mental Mahasiswa Angkatan 2022-2025

No	Q1 (UF)	Q2 (UF)	Q3 (UF)	Q4 (UF)	Q5 (UF)	Q6 (UF)	Q7 (UF)	Q8 (UF)	...	Q38 (F)
1	2	2	4	2	4	5	2	2	...	2
2	2	4	4	2	4	2	4	4	...	3
3	2	4	4	2	2	4	4	4	...	5
4	1	1	1	1	1	1	1	1	...	5
5	2	2	2	1	1	2	2	2	...	3
...
518	3	5	4	3	5	4	5	4	...	2
519	4	4	5	5	5	5	4	5	...	3
520	4	3	5	3	4	4	3	4	...	4
521	4	4	5	5	4	4	4	5	...	3
522	4	4	5	5	5	5	5	5	...	3

3.5 Implementasi K-Means Clustering

Algoritma K-Means diterapkan pada dataset yang telah melalui proses pembersihan dan transformasi, dengan total 522 data responden dan 38 atribut pertanyaan (Q1–Q38). Proses clustering dilakukan dengan variasi jumlah klaster ($k = 2$ hingga $k = 10$) menggunakan perhitungan jarak Euclidean dan Manhattan. Konfigurasi $k = 2$ dipilih untuk pembahasan karena memberikan hasil evaluasi yang relatif lebih baik dibandingkan konfigurasi lainnya berdasarkan Davies-Bouldin Index (DBI) dan Silhouette Coefficient. Algoritma K-Means yang diterapkan menggunakan perhitungan jarak Euclidean dan Manhattan sebagai parameter kedekatan antar data. Pada setiap iterasi, algoritma menghitung jarak antara setiap data dengan centroid masing-masing klaster, kemudian mengalokasikan setiap data ke klaster dengan jarak terdekat. Proses ini diulang secara berulang hingga posisi centroid tidak lagi mengalami perubahan yang signifikan, menandakan bahwa konvergensi telah tercapai[19].

Hasil pengelompokan disajikan pada Tabel 6, di mana setiap baris data telah diberi label klaster (Cluster_KMeans). Rekapitulasi jumlah data per klaster ditunjukkan pada Tabel 6, yakni Klaster 0 beranggotakan 334 mahasiswa (64,0%) dan Klaster 1 beranggotakan 188 mahasiswa (36,0%). Perbedaan jumlah anggota antar klaster menunjukkan adanya variasi distribusi responden berdasarkan hasil clustering. Namun, mengingat kualitas klaster yang masih rendah berdasarkan hasil evaluasi, pengelompokan ini lebih tepat dipandang sebagai analisis eksploratif daripada representasi kelompok yang terpisah secara jelas.

Tabel 5. Hasil Pengelompokan Cluster

No	Q1 (UF)	Q2 (UF)	Q3 (UF)	Q4 (UF)	Q5 (UF)	Q6 (UF)	Q7 (UF)	Q8 (UF)	...	Q38 (F)	Cluster_ KMeans
1	2	2	4	2	4	5	2	2	...	2	1
2	2	4	4	2	4	2	4	4	...	3	1
3	2	4	4	2	2	4	4	4	...	5	0
4	1	1	1	1	1	1	1	1	...	5	1



No	Q1 (UF)	Q2 (UF)	Q3 (UF)	Q4 (UF)	Q5 (UF)	Q6 (UF)	Q7 (UF)	Q8 (UF)	...	Q38 (F)	Cluster_ KMeans
5	2	2	2	1	1	2	2	2	...	3	1
...
518	3	5	4	3	5	4	5	4	...	2	0
519	4	4	5	5	5	5	4	5	...	3	0
520	4	3	5	3	4	4	3	4	...	4	0
521	4	4	5	5	4	4	4	5	...	3	0
522	4	4	5	5	5	5	5	5	...	3	0

Tabel 6. Jumlah Data setiap Cluster

Cluster	Jumlah Data
0	334
1	188

Dari hasil menggunakan algoritma K-Means berhasil mengelompokkan menjadi dua kategori utama berdasarkan pengelompokan kesehatan mental mahasiswa. Cluster 0 yang berjumlah 334 yang tergolong kesejahteraan psikologis, dan cluster 1 mendapatkan jumlah 188 yang tergolong tekanan psikologis. Hasil cluster ini berdasarkan rata-rata keseluruhan skor Q1-Q38, Cluster 0 memperoleh nilai sebesar 3,82, lebih tinggi dibandingkan cluster 1 sebesar 2,98. Berdasarkan rata-rata skor MHI-38, Klaster 0 memiliki rata-rata skor lebih tinggi dibandingkan Klaster 1. Penamaan klaster sebagai 'kesejahteraan psikologis' dan 'tekanan psikologis' hanya merupakan interpretasi deskriptif terhadap karakteristik skor, bukan klasifikasi psikologis yang bersifat definitif. Untuk memberikan gambaran yang lebih menyeluruh, dilakukan analisis silang antara hasil pembentukan klaster dengan atribut demografis responden, yang mencakup angkatan dan program studi. Sebaran mahasiswa di setiap klaster berdasarkan angkatan dapat diamati pada Tabel 7 berikut.

Tabel 7. Jumlah Cluster setiap Angkatan

Angkatan	Cluster 0	Cluster 1
2022	113	43
2023	73	50
2024	74	51
2025	74	44

Berdasarkan Tabel 7, distribusi mahasiswa ke dalam klaster berdasarkan angkatan menunjukkan bahwa Klaster 0 (kesejahteraan psikologis) mendominasi pada seluruh angkatan. Angkatan 2022 mencatat proporsi tertinggi mahasiswa dalam Klaster 0 dengan 113 mahasiswa (72,4%) dibandingkan Klaster 1 sebesar 43 mahasiswa (27,6%), mengindikasikan bahwa mahasiswa tingkat akhir cenderung berada dalam kondisi psikologis yang lebih stabil. Sementara itu, angkatan 2023 dan 2024 memperlihatkan perbandingan yang lebih berimbang dengan proporsi Klaster 1 mencapai sekitar 40%, mengindikasikan bahwa mahasiswa pada pertengahan masa studi menghadapi tekanan psikologis yang relatif lebih tinggi dibandingkan angkatan lainnya.

Tabel 8. Jumlah Cluster setiap Program Studi

Program Studi	Cluster 0	Cluster 1
Teknik Informatika	39	63
Teknik Industri	47	53
Sistem Informasi	66	34
Matematika	100	20
Teknik Elektro	82	18

Berdasarkan Tabel 8, distribusi klaster berdasarkan program studi memperlihatkan pola yang bervariasi antarprogram studi. Program studi Teknik Informatika merupakan satu-satunya program studi yang didominasi Klaster 1 (tekanan psikologis) dengan proporsi 61,8% (63 mahasiswa), diikuti Teknik Industri dengan proporsi Klaster 1 sebesar 53% (53 mahasiswa). Sebaliknya, program studi Matematika dan Teknik Elektro menunjukkan dominasi Klaster 0 paling tinggi, masing-masing sebesar 83,3% dan 82%, yang mengindikasikan bahwa mahasiswa pada kedua program studi tersebut cenderung memiliki kesejahteraan psikologis yang lebih baik dibandingkan program studi lain di Fakultas Sains dan Teknologi.

3.6 Pengujian

Setelah algoritma K-Means berhasil diimplementasikan, tahap berikutnya adalah melakukan evaluasi untuk mengukur kualitas hasil pengelompokan yang telah terbentuk. Proses evaluasi ini menggunakan sejumlah metrik yang lazim dipakai dalam analisis clustering, di antaranya Davies-Bouldin Index (DBI) dan Silhouette Coefficient.

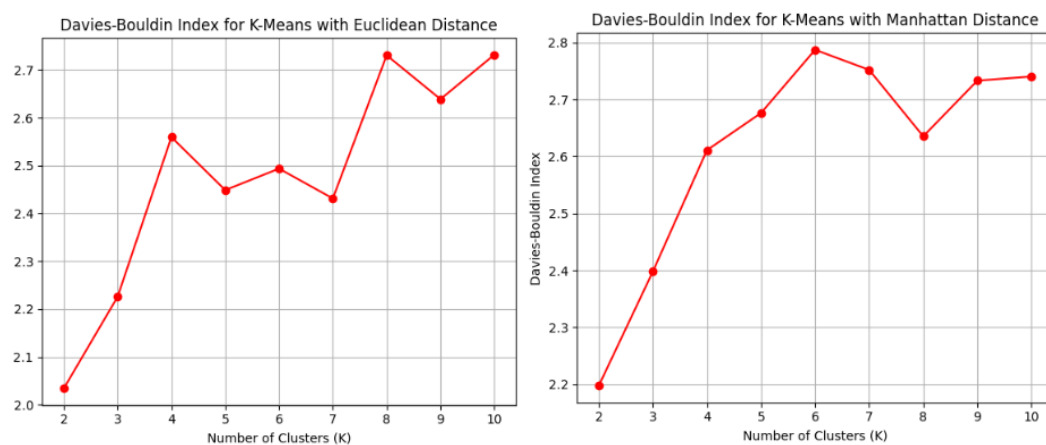
a. Davies-Bouldin Index (DBI)

DBI digunakan untuk mengevaluasi performa algoritma clustering dengan mengukur rasio antara jarak di dalam kluster terhadap jarak antar kluster. Nilai skor DBI yang semakin kecil atau mendekati nol menandakan bahwa algoritma tersebut semakin efektif dan optimal dalam mengelompokkan data[19]. Hal ini bertujuan untuk memastikan bahwa kluster yang terbentuk memiliki tingkat kepadatan internal yang tinggi (jarak antar-data di dalam satu kluster minimum) sekaligus memiliki jarak pemisahan yang maksimum antar-kluster, sehingga kluster terpisah secara tegas dan tidak tumpang tindih[20]. Berikut hasil pengujian menggunakan DBI:

Tabel 9. Hasil Pengujian DBI

Cluster	Hasil Euclidean	Hasil Manhattan
2	2.03	2.19
3	2.23	2.39
4	2.46	2.61
5	2.46	2.67
6	2.59	2.78
7	2.65	2.75
8	2.84	2.63
9	2.76	2.73
10	2.71	2.74

Berdasarkan hasil pengujian pada Tabel 9 dengan perhitungan jarak menggunakan *Euclidean Distance*, nilai DBI terendah diperoleh pada $k = 2$ dengan nilai 2,03. Nilai DBI meningkat seiring bertambahnya jumlah kluster. Hal ini menunjukkan bahwa penambahan jumlah kluster justru menurunkan kualitas pemisahan antar kelompok, sehingga $k = 2$ merupakan konfigurasi optimal berdasarkan metrik DBI. Dalam perhitungan jarak menggunakan *Manhattan Distance* dengan DBI, nilai terendah diperoleh pada $k=2$ dengan nilai 2,19.



Gambar 3. Visualisasi Hasil Pengujian DBI dengan Euclidean dan Manhattan

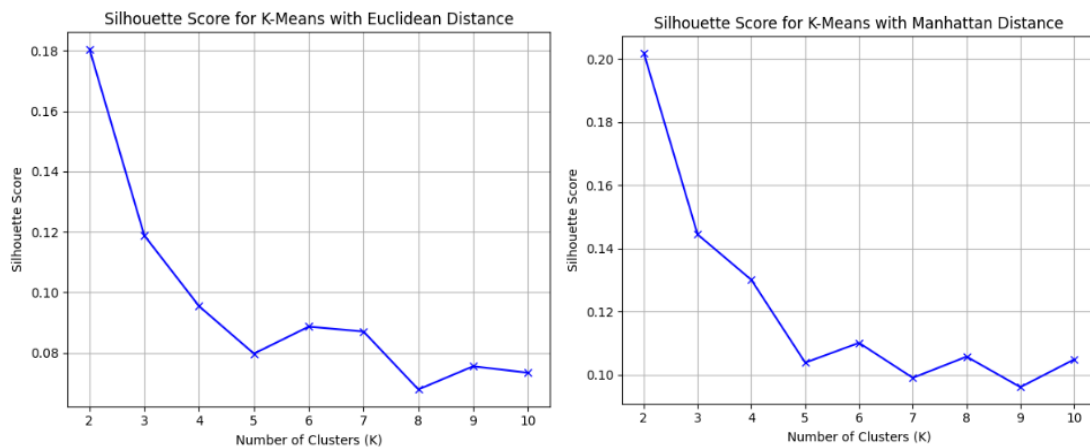
b. Silhouette Coefficient

Silhouette merupakan metode untuk interpretasi dan validasi konsistensi di dalam kluster data. Nilai silhouette mengukur seberapa mirip suatu objek dengan klasternya sendiri (kohesi) dibandingkan dengan kluster lain (separasi), dengan rentang nilai antara -1 hingga $+1$. Sebagaimana dijelaskan oleh[20], semakin tinggi skor Silhouette yang diperoleh, maka semakin kokoh dan valid kualitas pengelompokan yang dihasilkan oleh model clustering tersebut. Nilai tinggi ini menjadi pembuktian bahwa rata-rata jarak objek ke anggota kelompoknya sendiri jauh lebih kecil dibandingkan jaraknya ke kelompok lain, yang berarti proses pemisahan (separation) antar-kelompok sudah sangat ideal. Berikut hasil pengujian menggunakan Silhouette Coefficient:

Tabel 10. Hasil Pengujian Silhouette Coefficient

Cluster	Hasil Euclidean	Hasil Manhattan
2	0.18	0.20
3	0.13	0.14
4	0.11	0.13
5	0.10	0.10
6	0.08	0.11
7	0.08	0.09
8	0.06	0.10
9	0.06	0.09
10	0.06	0.10

Berdasarkan hasil pengujian pada Tabel 10, nilai Silhouette tertinggi menggunakan perhitungan jarak *Euclidean Distance* diperoleh pada $k = 2$ dengan nilai 0,18. Nilai ini menurun secara konsisten seiring bertambahnya jumlah kluster. Meskipun nilai 0,18 tergolong dalam kategori struktur kluster yang lemah, nilai ini tetap merupakan yang terbaik di antara seluruh konfigurasi yang diuji, sehingga memperkuat kesimpulan bahwa $k = 2$ adalah jumlah kluster yang paling optimal untuk dataset penelitian ini. Sedangkan dalam perhitungan jarak menggunakan *Manhattan Distance* didapatkan nilai tertinggi yaitu 0,20 pada $k=2$.



Gambar 4. Visualisasi Hasil Pengujian *Silhouette Coefficient* dengan Euclidean dan Manhattan

Analisis menggunakan Davies-Bouldin Index (DBI) dan Silhouette Coefficient menunjukkan bahwa pembagian data yang paling optimal adalah dua kluster. Hasil evaluasi menggunakan Davies-Bouldin Index (DBI) dan Silhouette Coefficient menunjukkan bahwa konfigurasi $k = 2$ memberikan nilai evaluasi terbaik dibandingkan konfigurasi lainnya yang diuji. Namun demikian, nilai Silhouette sebesar 0,18 (Euclidean) dan 0,20 (Manhattan) menunjukkan bahwa struktur kluster yang terbentuk masih lemah dan terdapat tingkat tumpang tindih antar kluster yang cukup tinggi. Oleh karena itu, hasil pengelompokan ini lebih tepat dipandang sebagai analisis eksploratif dan belum dapat dijadikan dasar yang kuat untuk pengambilan keputusan. Selain itu, analisis berdasarkan angkatan dan program studi hanya bersifat deskriptif karena tidak dilakukan uji signifikansi statistik, seperti Chi-Square. Penelitian selanjutnya disarankan untuk menerapkan metode reduksi dimensi, menggunakan algoritma clustering lain yang lebih sesuai, serta melakukan pengujian statistik inferensial agar hasil yang diperoleh memiliki validitas yang lebih kuat.

3.7 Pembahasan

Temuan penelitian ini menunjukkan keselarasan sekaligus perbedaan yang berarti dibandingkan penelitian terdahulu. Sun (2022) hanya menawarkan kerangka konseptual penerapan data mining dalam pendidikan kesehatan mental tanpa implementasi klusterisasi berbasis instrumen tervalidasi, sehingga tidak menghasilkan profil kluster yang dapat digunakan secara operasional [8]. Penelitian ini melengkapi celah tersebut dengan mengimplementasikan K-Means Clustering berbasis MHI-38 yang telah diadaptasi ke Bahasa Indonesia dan divalidasi psikolog ahli, menghasilkan dua kluster yang bermakna secara psikologis: Kluster 0 (kesejahteraan psikologis, 64%) dan Kluster 1 (tekanan psikologis, 36%).

Van Solang dan Nugroho (2023) mengklasifikasikan kesehatan mental mahasiswa ke dalam tiga kategori menggunakan K-Means tanpa instrumen psikometrik terstandar dan tanpa analisis demografis [9]. Penelitian ini berbeda dalam dua hal: penentuan jumlah kluster optimal ($k=2$) dilakukan secara kuantitatif melalui DBI dan Silhouette Coefficient, serta analisis silang berdasarkan angkatan dan program studi menghasilkan temuan yang lebih spesifik dan dapat ditindaklanjuti secara kebijakan. Sari et al. (2025) juga menggunakan K-Means pada mahasiswa STEM dengan evaluasi Silhouette, namun belum mengadaptasi instrumen secara kultural dan tidak membandingkan dua metode jarak secara simultan [10]. Sementara Alluri et al. (2026) lebih menekankan aspek keadilan pengelompokan (fairness-aware clustering) daripada pemetaan kondisi psikologis berbasis survei tervalidasi, sehingga berbeda secara substansial dari tujuan penelitian ini [11]. Secara keseluruhan, penelitian ini mengisi celah yang belum ditangani penelitian terdahulu melalui kombinasi instrumen MHI-38 yang diadaptasi secara kultural, perbandingan dua metode jarak, dan analisis demografis lintas angkatan serta program studi dalam satu institusi.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, algoritma K-Means Clustering dapat digunakan untuk mengidentifikasi pola kesehatan mental mahasiswa Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau menggunakan instrumen Mental Health Inventory-38 (MHI-38). Hasil pengelompokan menunjukkan bahwa konfigurasi dua kluster memberikan performa terbaik dibandingkan konfigurasi lainnya sehingga mampu memisahkan kelompok mahasiswa berdasarkan kecenderungan kesejahteraan psikologis dan tekanan psikologis. Meskipun kualitas kluster yang dihasilkan masih tergolong lemah berdasarkan evaluasi internal, hasil penelitian ini tetap memberikan gambaran awal mengenai distribusi

kondisi kesehatan mental mahasiswa dan dapat dimanfaatkan sebagai dasar penyusunan program pendampingan serta intervensi psikologis di lingkungan fakultas. Proses penelitian meliputi seleksi data, preprocessing, transformasi data, implementasi K-Means, serta evaluasi hasil clustering menggunakan Davies-Bouldin Index (DBI) dan Silhouette Coefficient pada variasi jumlah kluster $k = 2$ hingga $k = 10$. Hasil evaluasi menunjukkan bahwa konfigurasi $k = 2$ memberikan nilai evaluasi yang relatif lebih baik dibandingkan konfigurasi lainnya, dengan nilai DBI sebesar 2,03 dan Silhouette Coefficient sebesar 0,18 menggunakan Euclidean Distance, serta DBI sebesar 2,19 dan Silhouette Coefficient sebesar 0,20 menggunakan Manhattan Distance. Namun demikian, nilai Silhouette Coefficient yang masih berada di bawah 0,25 menunjukkan bahwa struktur kluster yang terbentuk masih lemah dan memiliki tingkat overlapping yang cukup tinggi, sehingga hasil pengelompokan ini lebih tepat dipandang sebagai analisis eksploratif dan belum dapat digunakan sebagai dasar yang kuat dalam pengambilan keputusan. Pengelompokan menghasilkan dua kluster dengan jumlah anggota 334 mahasiswa pada Kluster 0 dan 188 mahasiswa pada Kluster 1. Perbedaan karakteristik kedua kluster didasarkan pada kecenderungan rata-rata skor MHI-38. Analisis distribusi berdasarkan angkatan dan program studi menunjukkan adanya variasi proporsi antar kelompok, namun hasil tersebut hanya bersifat deskriptif karena penelitian ini tidak melakukan uji signifikansi statistik untuk memastikan bahwa perbedaan yang diamati benar-benar bermakna. Penelitian selanjutnya disarankan menggunakan metode clustering lain atau pendekatan reduksi dimensi agar kualitas pemisahan kluster dapat ditingkatkan. Penelitian ini memberikan kontribusi berupa pemetaan awal kondisi kesehatan mental mahasiswa Fakultas Sains dan Teknologi UIN Sultan Syarif Kasim Riau menggunakan algoritma K-Means Clustering berdasarkan instrumen Mental Health Inventory-38 (MHI-38).

REFERENCES

- [1] World Health Organization, *World Mental Health Report: Transforming Mental Health for All*. 2023. [Online]. Available: <https://www.who.int/publications/i/item/9789240055109>
- [2] N. Khoirunnisha and A. Sriati, "Stigma Kesehatan Mental dan Perilaku Mencari Bantuan Di Kalangan Mahasiswa: A Scoping Review," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 4, pp. 228–240, 2025, doi: 10.31004/riggs.v4i4.3366.
- [3] H. Wiyono, R. D. Ariyanto, N. S. Beda, M. S. Nuwa, and I. G. Juanamasta, "Masalah Psikososial pada Mahasiswa/I di Era Pandemi Covid-19," *J. Keperawatan Jiwa*, vol. 11, no. 1, p. 155, 2023, doi: 10.26714/jkj.11.1.2023.155-162.
- [4] L. Chen, H. Li, and Y. Zhang, "Academic Stress as a Predictor of Mental Health in University Students," *Cogent Educ.*, vol. 43, pp. 2893–2907, 2022, doi: 10.1007/s12144-022-03300-2.
- [5] W. A. Statistika, S. Y. U. Sihombing, R. K. Lubis, R. R. A. Situmorang, P. M. Fadilah, and S. Manullang, "Survei Kesehatan Mental Mahasiswa Menjelang Ujian Akhir Semester: Analisis Stres, Kecemasan, dan Strategi Penanggulangan," *J. Pendidik. Tambusai*, vol. 9(2), pp. 22722–22729, 2025.
- [6] Y. Astuti, I. R. Wulandari, A. R. Putra, and N. Kharomadhona, "Naïve Bayes untuk Prediksi Tingkat Pemahaman Kuliah Online Terhadap Mata Kuliah Algoritma Struktur Data," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 1, p. 28, 2022, doi: 10.26418/jp.v8i1.48848.
- [7] C. T. Veit and J. E. Ware, "The Structure of Psychological Distress and Well-Being in General Populations," *J. Consult. Clin. Psychol.*, vol. 51, no. 5, pp. 730–742, 1983, doi: 10.1037/0022-006X.51.5.730.
- [8] X. Sun, "Application of Data Mining Technology in College Mental Health Education," *Front. Psychol.*, vol. 13, no. August, pp. 1–10, 2022, doi: 10.3389/fpsyg.2022.974576.
- [9] T. G. Van Solang and A. Nugroho, "Analisis Kesehatan Mental Mahasiswa Menggunakan K-Means," *J. TEKINKOM*, vol. 6, no. 1, pp. 8–15, 2023, doi: 10.37600/tekinkom.v6i1.641.
- [10] F. Sari, M. Kuntari, W. Yati, H. Khaulasari, and M. Hafiyusholeh, "Implementasi K-Means Clustering Melalui Pemanfaatan Sampling Kombinasi Pada Pengelompokan Pola Kesehatan Mental Mahasiswa Sains dan Teknologi," *J. Nas. Teknol. dan Sist. Inf.*, vol. 11, no. 1, pp. 9–16, 2025, doi: 10.25077/teknosi.v11i01.2025.9-16.
- [11] P. Alluri, Z. Chen, T. Thesen, N. C. Jacobson, and W. J. Marrero, "Fairness-aware K-means Clustering in Digital Mental Health for Higher Education Students : a generalizable framework for equitable clustering," *JAMIA Open*, vol. 9, no. 1, pp. 1–11, 2026, doi: 10.1093/jamiaopen/ooaf174.
- [12] G. Bondolfi and P. Ghisletta, "Psychological Distress and Well-being among Students of Health Disciplines in Geneva , Switzerland : The Importance of Academic Satisfaction in The Context of Academic year- end and COVID-19 Stress on Their Learning Experience," *PLoS One*, vol. 433, no. June 2020, pp. 1–13, 2022, doi: 10.1371/journal.pone.0266612.
- [13] J. Wan, C. S. S. and , Lei Hum Wee, and Y. H. Wong, "Psychological Well-being and its Associated Factors Among University Students in Sichuan ," *Front. Psychol.*, vol. 16, no. February, pp. 1–11, 2025, doi: 10.3389/fpsyg.2025.1473871.
- [14] M. Muhamad Rizki, "Penerapan Metode K-Means Clustering Pada Data Penjualan," *J. Mhs. Tek. Inform.*, vol. 7, no. 2, pp. 1303–1307, 2023, doi: 10.36040/jati.v7i2.6562.
- [15] D. A. Muhammad Sholeh, Suraya, "Penerapan Data Mining pada Model Clustering Data Kuesioner Mahasiswa terhadap Kinerja Dosen," *EKSPLORA Inform.*, vol. 13, no. 2, pp. 208–217, 2023, doi: 10.30864/eksplora.v13i2.751.
- [16] M. Cui, "Introduction to the K-Means Clustering Algorithm Based on the Elbow Method," *Clausius Sci. Press*, vol. 1, no. 1, pp.



5–8, 2020, doi: 10.23977/accaf.2020.010102.

- [17] B. Chong, “K-means Clustering Algorithm : a brief review,” *Comput. Inf. Sci.*, vol. 4, no. 5, pp. 37–40, 2021, doi: 10.25236/AJCIS.2021.040506.
- [18] J. Lei, “An Analytical Model of College Students ’ Mental Health Education Based on the Clustering Algorithm,” *Math. Probl. Eng.*, vol. 2022, p. 11, 2022, doi: 10.1155/2022/1880214.
- [19] I. G. A. Gunadi and I. M. A. Wirawan, “Studi Perbandingan Algoritma Euclidean, Manhattan dan Chebysev Distance Untuk Optimasi Metode K-Means Clustering Pada Pengelompokan,” *Pendidik. Teknol. dan Kejuru.*, vol. 22, no. 2, pp. 116–127, 2025, doi: 10.33292/jptk.v22i2.54463.
- [20] M. Orisa and A. Faisol, “Analisis Algoritma Partitioning Around Medoid untuk Penentuan Klasterisasi,” *Teknol. Inf. dan Terap.*, vol. 8, no. 2, pp. 86–90, 2021, doi: 10/25047/jtit.v8i2.258.

