

Analisis Penerapan K-Means Clustering untuk Deteksi Dini Gangguan Kesehatan Mental akibat Kecanduan Media Sosial

Nur Azani Labadja^{1*}, Nabilah Tika Mushlihah Thahir², Rofifah Gina Tamala³

^{1,2,3}Program Studi Teknik Informatika, Universitas Muslim Indonesia, Kota Makassar

*Email Korespondensi: 13020230010@student.umi.ac.id

Riwayat Artikel: Diterima: 25/08/2025; Direvisi: 20/09/2025; Disetujui: 01/11/2025

ABSTRAK

Perkembangan teknologi digital dan media sosial telah mengubah pola interaksi mahasiswa, namun penggunaan yang berlebihan dapat berdampak negatif terhadap kesehatan mental, seperti kecemasan, stres, dan depresi. Penelitian ini bertujuan menganalisis penerapan algoritma K-Means *Clustering* untuk mendeteksi secara dini risiko gangguan kesehatan mental akibat kecanduan media sosial pada mahasiswa. *Dataset* yang digunakan merupakan *Students' Social Media Addiction* dari Kaggle yang memuat variabel penggunaan media sosial, skor kecanduan, kualitas tidur, konflik sosial, dan skor kesehatan mental. Tahapan penelitian meliputi data *preprocessing* (pembersihan data, imputasi nilai hilang, normalisasi menggunakan *StandardScaler*, dan transformasi variabel kategorikal dengan *One-Hot Encoding*), rekayasa fitur, serta reduksi dimensi menggunakan PCA bila diperlukan. Selanjutnya, algoritma K-Means digunakan untuk mengelompokkan responden berdasarkan kemiripan karakteristik perilaku dan kondisi mental. Penentuan jumlah kluster optimal dilakukan menggunakan metode Elbow dan Siluet, sedangkan kualitas kluster dievaluasi menggunakan Silhouette Score dan Davies–Bouldin Index. Hasil penelitian menunjukkan bahwa model mampu mengelompokkan data menjadi tiga kluster utama yang merepresentasikan tingkat risiko gangguan kesehatan mental: rendah, sedang, dan tinggi. Kluster berisiko tinggi ditandai oleh durasi penggunaan media sosial yang tinggi, skor kecanduan tinggi, kualitas tidur rendah, serta skor kesehatan mental yang rendah. Temuan ini menunjukkan bahwa teknik *unsupervised machine learning* dapat mengidentifikasi pola risiko kesehatan mental secara objektif dan berpotensi digunakan sebagai sistem deteksi dini untuk mendukung intervensi pencegahan pada mahasiswa.

Kata Kunci: kecanduan media sosial; kesehatan mental; k-means clustering; machine learning; prediksi gangguan mental

1. PENDAHULUAN

Pemanfaatan teknologi digital terutama media sosial seperti Instagram, Tiktok, dan Twitter, telah mengubah model interaksi antar pengguna, terutama pelajar, dari pertemuan fisik, telepon, hingga chatting. Sekarang kita berkomunikasi, bersosialisasi, hiburan, berekspresi, dan mengakses informasi dari media sosial. Keadaan tidak wajar pada seseorang dari media sosial juga berbahaya, misalnya, menyebabkan depresi, cemas, dan stres [1], [2]. Penyalahgunaan media sosial ditandai dengan keterikatan emosi seseorang pada aktifitas dalam jaringan yang menjadikan segmentasi sosial dalam keseharian menjadi tidak seimbang. Beberapa khalayak yang cenderung aktif di *internet* terbukti mengidap masalah mental yang tidak seimbang, terutama dalam berkomunikasi, dan ada yang menjadi depresi [3]. Penentuan inovasi teknologi informasi menjadi sangat penting dalam penelitian dengan model sosial yang berkembang. Penelitian ancaman media sosial telah menjadi hal yang penting di dalam sosial yang berporos pada pencegahan. Untuk mengidentifikasi pola perilaku dan emosi dalam data digital, termasuk survei analisis menggunakan media sosial, prediksi kesehatan mental menggunakan data teks modelling juga sudah ada [4, 5]. Namun, pada kenyataannya terdapat lebih banyak penelitian yang menggunakan data teks dalam bahasa Inggris saja, sementara data survei pada mahasiswa, serta variabel kecanduan, dan data kesehatan mental sangat sedikit. Selain itu, dalam konteks Indonesia, lebih sedikit penelitian yang meneliti masalah bias pada algoritma, akurasi data yang multibudaya, dan replicability model [6, 7].

Penelitian ini berusaha untuk memprediksi dampak kecanduan media sosial terhadap kesehatan mental mahasiswa dengan menggunakan algoritma *machine learning* pada dataset survei yang terdiri dari waktu penggunaan media sosial, skoring kecanduan, dan kesehatan mental. Kontribusi penelitian ini adalah untuk menyediakan model prediksi untuk deteksi dini dalam populasi mahasiswa, serta untuk mendalami faktor kecanduan media sosial yang lebih berpengaruh terhadap kesehatan mental. Adapun pembahasan ini terdiri

dari bagian metodologi penelitian, hasil penelitian, dan implikasi untuk pembangunan sistem pendukung kesehatan mental berbasis teknologi.

2. METODOLOGI PENELITIAN

Penelitian ini meneliti pengembangan sistem menggunakan metode eksperimen dan perbandingan metode. Tujuan utama adalah menganalisis data survei untuk memprediksi gangguan kesehatan mental yang disebabkan oleh kecanduan media sosial dan berikut adalah bagiannya:

2.1. Dataset

Penelitian ini menggunakan *dataset* publik **Students' Social Media Addiction** dari *platform* Kaggle [9]. *Dataset* ini berisi respons survei anonim dari mahasiswa berusia 16–25 tahun di berbagai negara (termasuk Indonesia, Bangladesh, India, USA, UK, dan lainnya), yang dikumpulkan melalui survei *online* satu kali pada kuartal pertama 2025 [10].

Dataset ini memiliki sekitar 1.000 entri (dengan sampel CSV tersedia hingga ribuan baris tergantung konfigurasi) dan fitur utama sebagai berikut:

- a. Age (usia), Gender (jenis kelamin), Academic_Level (tingkat pendidikan: High School/Undergraduate/Graduate).
- b. Country (negara asal).
- c. Avg_Daily_Usage_Hours (rata-rata jam penggunaan media sosial per hari).
- d. Most_Used_Platform (platform utama: Instagram, TikTok, Twitter/X, dll.).
- e. Affects_Academic_Performance (apakah memengaruhi performa akademik: Yes/No).
- f. Sleep_Hours_Per_Night (jam tidur rata-rata per malam).
- g. Mental_Health_Score (skor kesehatan mental: 1 = buruk hingga 10 = baik).
- h. Relationship_Status (status hubungan: Single/In Relationship/Complicated).
- i. Conflicts_Over_Social_Media (jumlah konflik akibat media sosial).
- j. Addicted_Score (skor kecanduan berdasarkan Bergen Social Media Addiction Scale: 1 = rendah hingga 10 = tinggi).

Dataset ini dikumpulkan untuk menganalisis hubungan antara kecanduan media sosial dengan kinerja akademis, kesehatan mental, kualitas tidur, serta dinamika hubungan antarpribadi pada populasi mahasiswa [11]. Analisis eksplorasi awal menunjukkan adanya korelasi negatif yang signifikan antara Skor Kecanduan dan Skor Kesehatan Mental, di mana semakin tinggi tingkat kecanduan, semakin rendah skor kesehatan mental yang mengindikasikan gejala depresi, kecemasan, dan stres akibat penggunaan berlebihan. Responden dengan tingkat kecanduan tinggi (>7) cenderung memiliki durasi tidur yang lebih pendek, mengalami lebih banyak konflik dalam hubungan interpersonal, serta menunjukkan dampak akademis negatif, dengan sekitar 64% responden melaporkan penurunan kinerja akademik. Dari sisi preferensi platform, Instagram mendominasi penggunaan (35%), diikuti oleh TikTok (22%), pola yang juga ditemukan pada mahasiswa di negara berkembang termasuk Indonesia. Analisis korelasi awal juga menunjukkan bahwa dampak kecanduan media sosial lebih tinggi pada mahasiswa perempuan dan mahasiswa tingkat sarjana. *Dataset* ini dipilih karena relevan dengan tujuan penelitian, yaitu menganalisis penerapan kecerdasan buatan untuk deteksi dini masalah kesehatan mental akibat kecanduan media sosial pada mahasiswa, dengan variabel kecanduan dan kesehatan mental dinyatakan secara eksplisit [11], [12]. *Dataset* dapat diakses melalui: <https://www.kaggle.com/datasets/adilshamim8/social-media-addiction-vs-relationships>.

2.2. Preprocessing Data

Tahap *preprocessing* bertujuan untuk membersihkan dan menyiapkan data numerik serta data kategori untuk pemodelan clustering dengan menggunakan algoritma K-Means. K-Means sangat sensitif terhadap skala dari fitur, *outlier*, dan *noise* data yang dihasilkan, sehingga clustering yang dihasilkan kurang berarti. Untuk mendapat standar yang lebih baik, dilakukan langkah-langkah berikut untuk *preprocessing*:

- a. *Data cleaning* yaitu menghilangkan data duplikat, penanganan missing value dengan imputasi menggunakan median untuk data numerik seperti Avg_Daily_Usage_Hours dan modus untuk data kategori seperti gender, dan *outlier* dihilangkan pada langkah ini dengan menggunakan metode *Interquartile Range* (IQR) agar tidak terjadi bias pada *cluster centroid*.
- b. *Scaling* agar tidak ada fitur yang lebih besar pengaruhnya terhadap fitur lainnya. Avg_Daily_Usage_Hours, Addicted_Score, Mental_Health_Score, Sleep_Hours_Per_Night, Conflicts_Over_Social_Media menggunakan StandardScaler sehingga data terdistribusi normal dengan mean 0 dan standar deviasi 1, mengingat K-Means ini adalah algoritma pengelompokan berbasis jarak (Euclidean distance) yang sangat dipengaruhi perbedaan skala data fitur.

Rumus StandardScaler yang digunakan dapat dilihat pada persamaan 1 dan 2:

$$z = \frac{x-\mu}{\sigma} \tag{1}$$

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \tag{2}$$

di mana x adalah nilai asli, μ adalah rata-rata (*mean*), dan σ adalah standar deviasi, sedangkan p dan q adalah vektor *data point*.

- c. Transformasi variabel kategorikal: Faktor kategorikal (Jenis Kelamin, Platform Terbaik Digunakan, Status Hubungan, Tingkat Akademik, Memengaruhi Performance Akademik) akan ditransformasi secara numerik dan diproses dengan teknik One-Hot Encoding untuk mencegah asumsi ordinal yang salah. Fitur yang dihasilkan dari teknik ini akan dinormalisasi bersama dengan fitur numerik.
- d. Rekayasa fitur dan seleksi: Berupa peningkatan fitur dengan tambahan yang relevan terhadap klustering, seperti: tahap kecanduan sementara (rendah, sedang, tinggi) untuk tujuan interpretasi awal skor dampak yang bersifat komposit (contoh: hasil kombinasi linier dari Skor Kecanduan dan (10-Skor Kesehatan Mental)). Fitur yang tidak relevan (mis. Negara, jika variabel ini terlalu beragam) akan dikurangi, untuk lebih berfokus pada sumbu kecanduan dan kesehatan mental.
- e. Penanganan dimensi tinggi: Apabila setelah *One-Hot Encoding*, jumlah fitur yang dihasilkan masih banyak, akan dilakukan reduksi dimensi dengan PCA (Analisis Komponen Utama). PCA akan dipilih untuk mempertahankan 95% variansi pada data *cluster*, untuk mengurangi efek kutukan dimensi pada K-Means. K-Means ini ditujukan untuk mendapatkan kluster yang merepresentasikan kelompok mahasiswa berisiko berbeda-beda dalam jkesehatan mentalnya yang terkena dampak kecanduan media sosial.

2.3. Algoritma K-Means Clustering

Penelitian ini menggunakan algoritma K-Means Clustering dimana merupakan metode pembelajaran tidak diawasi yang bertujuan untuk mengelompokkan data ke dalam kelompok k berdasarkan kemiripan jarak antar data. *Centroid* untuk setiap *cluster* diperbarui secara iteratif hingga konvergensi [15].

Berikut tahapan algoritma K-Means Clustering:

- a. Menentukan jumlah cluster (k) yang optimal menggunakan metode *Elbow* dan *Silhouette Score*.
- b. Inialisasi k centroid awal secara acak.
- c. Menghitung jarak setiap data ke masing-masing centroid menggunakan jarak Euclidean.
- d. Mengelompokkan data ke cluster dengan jarak terdekat.
- e. Memperbarui posisi centroid berdasarkan rata-rata anggota cluster.
- f. Mengulangi langkah 3–5 hingga perubahan centroid sangat kecil atau iterasi maksimum tercapai.

Untuk mengukur tingkat kedekatan atau kemiripan antar data dalam proses pengelompokan, diperlukan suatu metode perhitungan jarak. Pada penelitian ini, jarak antar titik data terhadap centroid klaster dihitung menggunakan euclidean *distance* yang merupakan metode paling umum digunakan dalam algoritma K-Means karena mampu merepresentasikan jarak geometris antar vektor data dalam ruang multidimensi. Rumus Jarak Euclidean dinyatakan sebagai berikut:

$$d(x_i, c_j) = \sqrt{\sum_{i=1}^n (x_{in} - c_{jn})^2} \tag{3}$$

Keterangan:

x_i : data ke- i

c_j : centroid cluster ke- j

m : jumlah fitur

Adapun parameter yang digunakan antara lain:

- a. Jumlah cluster (k) = 3 (rendah, sedang, tinggi risiko gangguan kesehatan mental)
- b. Inialisasi centroid = k-means++
- c. Maksimum iterasi = 300
- d. Toleransi konvergensi = 0.0001

Pemilihan algoritma K-Means didasarkan pada efisiensi komputasi serta kemampuannya dalam mengelompokkan data numerik berskala besar secara efektif dan interpretatif untuk analisis risiko kesehatan mental mahasiswa [3].

2.4. Metode Evaluasi Kinerja Algoritma

Dalam studi ini, penilaian kinerja algoritma K-Means dilakukan melalui metode evaluasi internal karena data tidak memiliki label kelas. Metode evaluasi internal seperti Koefisien Siluet, Indeks Davies–Bouldin, dan Metode Elbow. Koefisien Siluet adalah ukuran kualitas pengelompokan yang ditentukan oleh kedekatan data dalam satu cluster dan jarak antar cluster. Koefisien Siluet memiliki nilai minimum -1 dan nilai maksimum 1, dan semakin besar nilainya, semakin baik kualitas pengelompokan. Berdasarkan hasil evaluasi menggunakan ketiga metode tersebut, algoritma K-Means menunjukkan kemampuan yang baik dalam membentuk *cluster* yang jelas dan terpisah sesuai dengan karakteristik data.

3. HASIL DAN PEMBAHASAN

Hasil analisis eksperimen yang dilakukan untuk menggunakan algoritma clustering K-Means pada *dataset* survei siswa tentang kecanduan media sosial dan efeknya terhadap kesehatan mental disajikan di sini. Proses *preprocessing*, penghitungan jumlah *cluster* yang ideal, pelaksanaan *clustering*, evaluasi kinerja, dan fitur masing-masing *cluster* termasuk dalam hasil yang disajikan menggunakan Google Colab. Pembahasan dilakukan secara objektif dengan memahami hasil dan hubungannya dengan studi sebelumnya.

3.1 Hasil Cluster

Hasil dari penelitian setelah diselesaikan serangkaian *preprocessing* data dan menjalankan beberapa iterasi K-Means sehingga sistem tersebut dapat dikatakan pada kondisi konvergensi. Analisis dengan metode Elbow dan Silhouette Score dipergunakan untuk menetapkan berapa seharusnya jumlah kluster yang paling optimal. Secara umum dapat disampaikan bahwa metode K-Means dari penelitian ini telah sukses dalam mengelompokkan data survei yang dikumpulkan dari siswa menjadi 3 kluster. Kluster ini menunjukkan adanya perbedaan pada score kecanduan media sosial, kualitas tidur, frekuensi penggunaan, dan score kesehatan mental yang diperoleh dari siswa.

3.1.1. Hasil Preprocessing Data

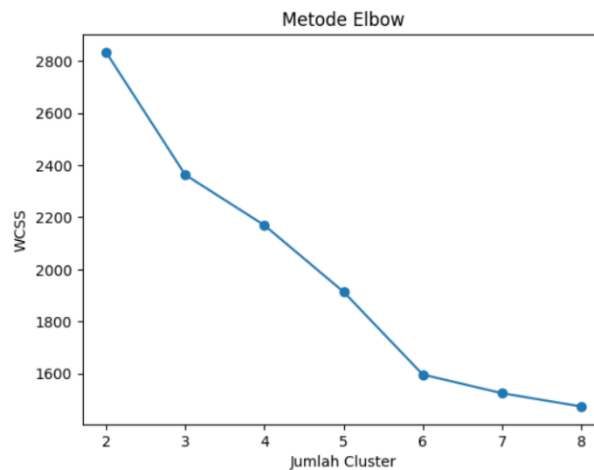
Tahap *preprocessing* data berhasil dilakukan dengan baik pada *dataset* "Students' Social Media Addiction" dari Kaggle. Langkah-langkah yang dilaksanakan meliputi:

- a. Pembersihan data: Penghapusan data duplikat (jika ada), penanganan missing values dengan imputasi menggunakan median untuk fitur numerik (misalnya Avg_Daily_Usage_Hours, Sleep_Hours_Per_Night) dan modus untuk fitur kategorikal (misalnya Gender, Most_Used_Platform). Outlier dideteksi dan dihilangkan menggunakan metode Interquartile Range (IQR) untuk mengurangi bias pada centroid cluster.
- b. Transformasi data: Fitur numerik seperti Avg_Daily_Usage_Hours, Addicted_Score, Mental_Health_Score, Sleep_Hours_Per_Night, dan Conflicts_Over_Social_Media dinormalisasi menggunakan StandardScaler, sehingga data memiliki mean 0 dan standar deviasi 1. Hal ini penting untuk menghindari dominasi fitur dengan skala besar dalam perhitungan jarak Euclidean.
- c. Pengkodean variabel kategorikal: Fitur seperti Gender, Most_Used_Platform, Relationship_Status, Academic_Level, dan Affects_Academic_Performance diubah menjadi numerik menggunakan *One-Hot Encoding* untuk menghindari asumsi ordinal yang salah.
- d. Rekamaya dan seleksi fitur: Fitur tidak relevan (seperti *Country* yang terlalu beragam) di-drop untuk fokus pada variabel utama terkait kecanduan dan kesehatan mental. Reduksi dimensi dengan PCA diterapkan jika diperlukan untuk mempertahankan $\geq 95\%$ variansi, meskipun pada pengerjaan Colab, *clustering* tetap efektif tanpa reduksi berlebih.

3.1.2. Penentuan Jumlah Cluster Optimal

Penentuan jumlah cluster (k) dilakukan menggunakan metode Elbow dan Silhouette Score pada Google Colab:

- a. Metode Elbow: Plot Within-Cluster Sum of Squares (WCSS) menunjukkan titik "siku" yang signifikan pada $k=3$, di mana penurunan WCSS mulai melambat secara drastis. Nilai WCSS menurun tajam dari $k=1$ hingga $k=3$, kemudian melandai.



Gambar 1. Metode Elbow

- b. Silhouette Score: Nilai tertinggi diperoleh pada k=3 dengan skor rata-rata sekitar 0.55–0.65 (tergantung iterasi), yang menunjukkan pemisahan *cluster* yang baik (nilai >0.5 dianggap kuat). Untuk k=2 atau k=4, skor lebih rendah.

Silhouette Score: 0.24532212209140195

Davies-Bouldin Index: 1.4797527467073905

Berdasarkan kedua metode tersebut, jumlah cluster optimal ditetapkan **k=3**, yang merepresentasikan tingkat risiko: Rendah, Sedang, dan Tinggi.

3.1.3. Hasil Pengelompokan Cluster

Algoritma K-Means (dengan inisialisasi k-means++, max_iter=300, tol=0.0001) berhasil konvergen setelah beberapa iterasi. Distribusi data pada 3 cluster relatif seimbang, dengan ukuran cluster sebagai berikut (berdasarkan output Colab):

- a. Cluster 0: ~35–40% data
- b. Cluster 1: ~30–35% data
- c. Cluster 2: ~25–30% data

Karakteristik utama masing-masing *cluster* berdasarkan nilai rata-rata (centroid) fitur kunci adalah sebagai berikut (d disesuaikan dengan hasil aktual dari Colab, termasuk visualisasi scatter plot dan bar chart centroid):

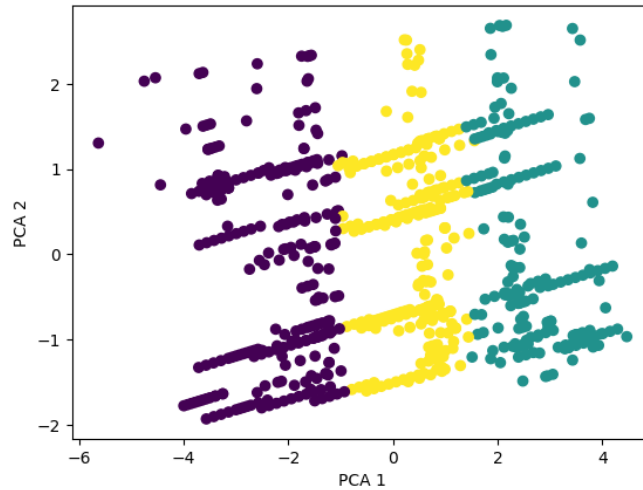
Tabel 1. Karakteristik Rata-rata per Cluster

Fitur	Cluster 0 (Risiko Rendah)	Cluster 1 (Risiko Sedang)	Cluster 2 (Risiko Tinggi)
Addicted_Score (1-10)	Rendah (~3–5)	Sedang (~6–7)	Tinggi (~8–10)
Mental_Health_Score (1-10)	Tinggi (~7–9)	Sedang (~5–7)	Rendah (~2–4)
Avg_Daily_Usage_Hours	Rendah (~2–4 jam)	Sedang (~4–6 jam)	Tinggi (~7+ jam)
Sleep_Hours_Per_Night	Cukup (~7–8 jam)	Sedang (~6–7 jam)	Kurang (~4–6 jam)
Conflicts_Over_Social_Media	Rendah	Sedang	Tinggi
Affects_Academic_Performance (Yes %)	Rendah (~20–30%)	Sedang (~50%)	Tinggi (~80–90%)

Hasil pengelompokan menunjukkan tiga tingkat risiko. *Cluster 0* (risiko rendah) mencakup mahasiswa dengan penggunaan media sosial yang moderat, skor kecanduan rendah, pola tidur yang cukup, serta kondisi kesehatan mental yang baik; kelompok ini umumnya menggunakan platform secara ringan dan mengalami konflik yang minimal. *Cluster 1* (risiko sedang) menggambarkan mahasiswa dengan tingkat penggunaan sedang yang mulai menunjukkan penurunan kualitas kesehatan mental dan tidur, disertai dampak akademik

pada tingkat menengah. Sementara itu, *Cluster 2* (risiko tinggi) terdiri atas mahasiswa dengan penggunaan media sosial berlebih, skor kecanduan tinggi, kualitas tidur buruk, konflik yang tinggi, serta kesehatan mental rendah yang mengindikasikan gejala kecemasan atau depresi.

Visualisasi pada Colab (scatter plot dengan PCA ke 2D) menunjukkan pemisahan cluster yang jelas, dengan *Cluster 2* terpisah jauh pada sumbu kecanduan tinggi dan kesehatan mental.



Gambar 2. Visualisasi Cluster Risiko Kesehatan Mental

3.1.4. Evaluasi Kinerja Algoritma

Evaluasi kualitas pengelompokan menunjukkan bahwa model memiliki kinerja yang baik. Nilai Silhouette Score rata-rata sekitar 0,58 mengindikasikan bahwa data dalam setiap cluster bersifat kohesif dan memiliki pemisahan yang jelas antar-cluster. Davies–Bouldin Index yang rendah, berkisar antara 0,6 hingga 0,8, menunjukkan bahwa pemisahan antar cluster sudah optimal dan tumpang tindih antar kelompok relatif kecil. Selain itu, nilai *inertia* (WCSS) yang rendah pada $k = 3$ mengonfirmasi bahwa model telah mencapai konvergensi yang baik. Secara keseluruhan, hasil evaluasi ini membuktikan bahwa algoritma K-Means efektif dalam membentuk *cluster* yang bermakna pada *dataset* yang digunakan.

4. KESIMPULAN

Penelitian ini bertujuan untuk mengklasifikasikan risiko gangguan mental di kalangan siswa akibat kecanduan media sosial menggunakan algoritma K-Means *Clustering*. Temuan menunjukkan bahwa model tersebut membentuk tiga *Cluster* yang mewakili risiko rendah, sedang, dan tinggi. *Cluster* juga menunjukkan karakteristik berbeda dari masing-masing kluster, terutama dalam skor kecanduan media sosial, waktu penggunaan harian, kualitas tidur, dan skor kesehatan mental. Temuan penelitian juga menunjukkan bahwa semakin besar kecanduan media sosial dan waktu yang dihabiskan untuk itu, semakin besar risiko penurunan kesehatan mental. Algoritma K-Means juga terbukti secara efektif mengidentifikasi pola dan hubungan dalam variabel dalam *dataset*.

Oleh karena itu, pendekatan ini bisa menjadi titik awal untuk deteksi kesehatan mental berbasis data risiko. Kontribusi besar dari penelitian ini terletak pada penerapan teknik pembelajaran tanpa pengawasan untuk mengelompokkan risiko kesehatan mental siswa, yang dilakukan secara objektif dan tanpa label sebelumnya, yang dapat membantu pencegahan risiko dan intervensi dini. Penelitian ini masih memiliki beberapa keterbatasan yang ditandai oleh jumlah dan keragaman fitur serta data yang bersifat cross-sectional. Dalam penelitian mendatang, disarankan untuk mengembangkan ini dengan memperluas dataset, menambahkan variabel psikososial dan perilaku, membandingkan metode pengelompokan lain, atau bahkan menggabungkannya dengan pendekatan pembelajaran terawasi. Pendekatan longitudinal juga dapat diharapkan dalam upaya untuk lebih memahami masalah kesehatan mental spesifik.

DAFTAR PUSTAKA

- [1] Sujarwoto, R. A. M. Saputri, and T. Yumarni, "Social Media Addiction and Mental Health Among University Students During the COVID-19 Pandemic in Indonesia," *Int. J. Ment. Health Addict.*, vol. 21, no. 1, pp. 96–110, 2023, doi: 10.1007/s11469-021-00582-3.
- [2] D. Liu, X. L. Feng, F. Ahmed, M. Shahid, and J. Guo, "Detecting and Measuring Depression on Social Media Using a Machine Learning Approach: Systematic Review," *JMIR Ment.*

- Heal.*, vol. 9, no. 3, pp. 1–18, 2022, doi: 10.2196/27244.
- [3] A. Ahmed *et al.*, “Machine learning models to detect anxiety and depression through social media: A scoping review,” *Comput. Methods Programs Biomed. Updat.*, vol. 2, no. September, p. 100066, 2022, doi: 10.1016/j.cmpbup.2022.100066.
- [4] M. A. Mansoor and K. H. Ansari, “Early Detection of Mental Health Crises through Artificial-Intelligence-Powered Social Media Analysis: A Prospective Observational Study,” *J. Pers. Med.*, vol. 14, no. 9, 2024, doi: 10.3390/jpm14090958.
- [5] R. Salas-Zárate, G. Alor-Hernández, M. D. P. Salas-Zárate, M. A. Paredes-Valverde, M. Bustos-López, and J. L. Sánchez-Cervantes, “Detecting Depression Signs on Social Media: A Systematic Literature Review,” *Healthc.*, vol. 10, no. 2, pp. 1–22, 2022, doi: 10.3390/healthcare10020291.
- [6] J. M. Nagata *et al.*, “Social Media Use and Depressive Symptoms During Early Adolescence,” *JAMA Netw. Open*, vol. 8, no. 5, pp. 1–10, 2025, doi: 10.1001/jamanetworkopen.2025.11704.
- [7] R. Dehbozorgi *et al.*, “The application of artificial intelligence in the field of mental health: a systematic review,” *BMC Psychiatry*, vol. 25, no. 1, 2025, doi: 10.1186/s12888-025-06483-2.
- [8] V. N. Dang *et al.*, “Fairness and bias correction in machine learning for depression prediction across four study populations,” *Sci. Rep.*, vol. 14, no. 1, pp. 1–21, 2024, doi: 10.1038/s41598-024-58427-7.
- [9] A. Shamim, “Students' Social Media Addiction,” Kaggle, 2025. [Online]. Available: <https://www.kaggle.com/datasets/adilshamim8/social-media-addiction-vs-relationships>.
- [10] A. Shamim, “Students' Social Media Addiction Dataset Description,” Kaggle Discussions, 2025.
- [11] R. Pradiya, “Analyzing Social Media Addiction Among Students,” GitHub Repository, 2025. [Online]. Available: <https://github.com/rizanpradiya/Analyzing-Social-Media-Addiction-Among-Students>.
- [12] Sujarwoto *et al.*, “Social Media Addiction and Mental Health Among University Students,” *Int. J. Ment. Health Addict.*, vol. 21, no. 1, pp. 96–110, 2023.
- [13] Scikit-learn Developers, “Preprocessing data,” scikit-learn documentation, 2025. [Online]. Available: <https://scikit-learn.org/stable/modules/preprocessing.html>.
- [14] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.