

SKRIPSI

**IMPLEMENTASI K-MEANS CLUSTERING UNTUK SEGMENTASI DAN
PROFILING MINAT BACA MAHASISWA UNIVERSITAS PALANGKA
RAYA**



DISUSUN OLEH

NAMA : MELINDA SEPTA NATALINA

NIM : 213020503085

JURUSAN TEKNIK INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS PALANGKA RAYA

2025/2026

**IMPLEMENTASI K-MEANS CLUSTERING UNTUK SEGMENTASI DAN PROFILING MINAT
BACA MAHASISWA UNIVERSITAS PALANGKA RAYA**

SKRIPSI

bagai salah satu syarat untuk menyelesaikan Program Strata-1 pada Jurusan Teknik Informatika Fakultas
Teknik Universitas Palangka Raya

Oleh

MELINDA SEPTA NATALINA

213020503085

Telah dipertahankan didepan tim penguji, pada :

Hari/Tanggal : Kamis, 16 Juli 2026

Waktu : 11:00-12:30 WIB

Tempat : Ruang Ujian 1

1. **NOVERA KRISTIANTI, S.T., M.T.**

NIP. 19931116 202203 2 012

..... (Ketua Penguji)

2. **RESSA PRISKILA, S.T., M.T.**

NIP. 19940301 201903 1 013

..... (Pembimbing Utama/Pertama)

3. **SEPTIAN GEGES, S.Kom., M.Kom.**

NIP. 19910917 202012 1 004

..... (Pembimbing Pendamping/Kedua)

4. **FELICIA SYLVIANA, S.T., M.M.**

NIP. 19760118 200312 2 003

..... (Anggota)

Mengetahui :

**Jurusan / Program Studi Teknik Informatika
Fakultas Teknik Universitas Palangka Raya
Ketua Jurusan,**

**ARIESTA LESTARI, S.Kom., M.Cs., Ph.D.
NIP. 19800322 200501 2 004**

HALAMAN PERNYATAAN

Dengan ini saya menyatakan dengan sebenar-benarnya bahwa dalam Skripsi ini tidak terdapat karya ilmiah yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, serta tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali secara tertulis diacu dalam Skripsi ini dan disebutkan dalam Tinjauan Pustaka serta Daftar Pustaka.

Palangka Raya, 21 Juli 2026



Melinda Septa Natalina

213020503085

BIODATA DIRI

Data Diri

Nama : Melinda Septa Natalina
NIM : 213020503085
Fakultas : Teknik
Jurusan/Program Studi : Teknik Informatika
Jenjang : Strata 1 (S-1)
Jenis Kelamin : Perempuan
Tempat, Tanggal Lahir : Sabailah, 23 September 2003
Agama : Kristen Protestan
Status Dalam Keluarga : Anak kandung
Anak Ke : 1 dari 2 bersaudara
Alamat : Jl. Pahlawan No.27 RT.32 RW.004 Buntok
No Telepon/HP : 082252842509



Data Orang Tua

Nama Ayah : Partaminato
Pekerjaan Ayah : Swasta
Nama Ibu : Rimaniati
Pekerjaan Ibu : Pegawai Negeri Sipil
Alamat Orang Tua : Jalan Pahlawan No.27 RT.32 RW.004 Buntok
No.HP Orang Tua : 081251533970

Riwayat Pendidikan

SD : SDS Santa Maria Buntok (Tahun Lulus 2015)
SMP : SMP Santa Maria Buntok (Tahun Lulus 2018)
SMA : SMAN 1 Dusun Selatan (Tahun Lulus 2021)

Palangka Raya, 21 Juli 2026

MELINDA SEPTA NATALINA

HALAMAN PERSEMBAHAN

Puji syukur yang tak terhingga penulis panjatkan kepada Tuhan Yang Maha Esa atas kasih, penyertaan, dan anugerah-Nya yang tidak pernah berhenti mengalir dalam setiap langkah kehidupan penulis. Skripsi ini penulis persembahkan :

1. Kepada Tuhan yang senantiasa menguatkan, memelihara, dan menyertai penulis, bahkan di saat-saat terendah dalam proses perjalanan ini. Firman Tuhan dalam Yesaya 41:10 senantiasa menjadi penguat bagi penulis: *"Janganlah takut, sebab Aku menyertai engkau, janganlah bimbang, sebab Aku ini Allahmu; Aku akan meneguhkan, bahkan akan menolong engkau; Aku akan memegang engkau dengan tangan kanan-Ku yang membawa kemenangan."*
2. Skripsi ini penulis persembahkan kepada kedua orang tua tercinta yang tidak pernah berhenti memberikan kasih sayang, dukungan, doa, pengorbanan, serta kerja keras yang luar biasa demi masa depan penulis. Terima kasih karena selalu menjadi tempat pulang, sumber kekuatan, dan alasan penulis untuk terus berjuang hingga mampu menyelesaikan tugas akhir ini. Kepada adik tercinta serta seluruh keluarga besar, terima kasih atas doa, dukungan, perhatian, dan semangat yang selalu diberikan kepada penulis.
3. Terima kasih kepada Ibu Ressa Priskila, S.T., M.T. selaku Dosen Pembimbing I sekaligus Dosen Pembimbing Akademik yang telah dengan sabar membimbing, mengarahkan, memberikan masukan, serta mendampingi penulis hingga seluruh proses penyusunan skripsi ini dapat terselesaikan dengan baik. Terima kasih pula kepada Dosen Pembimbing II Bapak Septian Geges, S.Kom., M.Kom yang telah memberikan ilmu, arahan, dan motivasi selama proses penelitian berlangsung. Kepada seluruh dosen Program Studi Teknik Informatika Universitas Palangka Raya, penulis mengucapkan terima kasih atas ilmu, pengalaman, dan pembelajaran berharga yang telah diberikan selama masa perkuliahan.
4. Skripsi ini penulis persembahkan kepada para sahabat yang selalu hadir dalam perjalanan hidup penulis, yaitu anggota KEMEH, CS, SWAGGIE, Sana Sana Sini, Belia, Bobby, serta semua sahabat yang tidak dapat

disebutkan satu per satu. Terima kasih karena telah menjadi bagian dari cerita, tempat berbagi tawa, tangis, keluh kesah, dan semangat selama perjalanan perkuliahan hingga penyelesaian skripsi ini. Meskipun tidak semua nama tertulis di halaman ini, kalian akan selalu hidup dalam hati dan ingatan penulis.

5. Terima kasih kepada keluarga besar Duta Bahasa, Meine Welt, dan Utama yang telah menjadi tempat penulis bertumbuh, belajar, berkembang, serta menemukan banyak pengalaman berharga. Terima kasih karena telah menjadi rumah kedua yang menghadirkan begitu banyak cerita, pembelajaran, dukungan, dan kenangan yang tidak akan pernah terlupakan.
6. Kepada **R**, terima kasih telah hadir dan menemani perjalanan penulis sejak Seminar Proposal hingga Sidang Akhir. Terima kasih atas waktu, dukungan, perhatian, kesabaran, dan kehadiran yang selalu memberikan kekuatan di setiap proses yang dilalui. Rasa terima kasih ini mungkin tidak akan pernah cukup untuk menggambarkan betapa berharganya pendampingan yang telah diberikan kepada penulis hingga saat ini.
7. Terima kasih kepada Taeyong dan seluruh anggota NCT yang secara tidak langsung telah menemani proses penyusunan skripsi ini melalui karya-karya, musik, semangat, dan berbagai hal sederhana yang mampu menghibur penulis di tengah rasa lelah dan tekanan selama menyelesaikan tugas akhir.
8. Kepada Rinanda Maharani dan Salma Ranggita, Puteri Indonesia Pendidikan 2025 dan Puteri Indonesia Pariwisata 2025, terima kasih karena telah menjadi sumber inspirasi bagi penulis. Dedikasi, kecerdasan, kepedulian, dan kontribusi yang kalian tunjukkan menjadi pengingat bahwa perempuan Indonesia memiliki kemampuan luar biasa untuk berkarya, menginspirasi, dan membawa perubahan positif bagi masyarakat.
9. Skripsi ini penulis persembahkan untuk diri penulis sendiri, Melinda Septa Natalina. Terima kasih karena telah bertahan, berjuang, dan tidak menyerah

meskipun harus melalui berbagai tantangan, kegagalan, keraguan, dan air mata dalam perjalanan ini. Terima kasih karena tetap melangkah ketika keadaan terasa berat, tetap bangkit ketika jatuh, dan tetap percaya bahwa semua usaha akan menemukan hasilnya pada waktu yang tepat. Perjalanan ini bukanlah perjalanan yang mudah, tetapi penulis berhasil sampai pada titik ini. Untuk setiap luka, pelajaran, proses, dan kemenangan kecil yang berhasil dilewati, penulis pantas berbangga. Semoga pencapaian ini menjadi awal dari perjalanan panjang untuk terus bertumbuh, berkarya, dan menjadi terang serta garam dunia bagi sesama.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas kasih karunia, penyertaan, dan berkat-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “**Implementasi K-Means Clustering untuk Segmentasi dan Profiling Minat Baca Mahasiswa Universitas Palangka Raya**” sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik pada Program Studi Teknik Informatika, Fakultas Teknik, Universitas Palangka Raya.

Dalam proses penyusunan skripsi ini, penulis memperoleh banyak bantuan, dukungan, bimbingan, serta doa dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada kedua orang tua dan keluarga tercinta atas segala doa dan dukungan yang diberikan, kepada dosen pembimbing serta seluruh dosen Program Studi Teknik Informatika Universitas Palangka Raya atas ilmu dan arahan yang diberikan, kepada sahabat, rekan seperjuangan, dan seluruh pihak yang telah membantu selama proses penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan menjadi referensi bagi penelitian selanjutnya.

Palangka Raya, 21 Juli 2026

Penulis,

Melinda Septa Natalina

213020503085

IMPLEMENTASI K-MEANS CLUSTERING UNTUK SEGMENTASI DAN PROFILING MINAT BACA MAHASISWA UNIVERSITAS PALANGKA RAYA

Melinda Septa Natalina (213020503085)

Jurusan Teknik Informatika Fakultas Teknik Universitas Palangka Raya

Kampus Tunjung Nyaho Jl. Yos Sudarso Palangka Raya 73112

Email : melindaseptaa23@gmail.com

ABSTRAK

Penelitian ini bertujuan mengelompokkan mahasiswa Universitas Palangka Raya berdasarkan karakteristik minat baca untuk mendukung perancangan program literasi dan layanan perpustakaan yang lebih tepat sasaran. Penelitian menggunakan metode *K-Means Clustering* yang dikombinasikan dengan *Principal Component Analysis (PCA)* pada data kuesioner dari 394 mahasiswa. Tahapan penelitian meliputi preprocessing data, normalisasi menggunakan *StandardScaler*, reduksi dimensi menggunakan PCA, penentuan jumlah cluster optimal dengan *Elbow Method*, serta evaluasi menggunakan *Silhouette Score* dan *Davies-Bouldin Index*. Hasil PCA mereduksi 40 fitur menjadi 16 komponen utama, sedangkan proses clustering menghasilkan empat cluster, yaitu Pembaca Aktif, Pembaca Selektif, Pembaca Kasual, dan Pembaca Pasif. Nilai *Silhouette Score* sebesar 0,2125 dan *Davies-Bouldin Index* sebesar 1,8785 menunjukkan kualitas cluster berada pada kategori moderat. Hasil segmentasi menunjukkan bahwa 61,2% mahasiswa berada pada kelompok Pembaca Kasual dan Pembaca Pasif dengan aktivitas membaca yang relatif rendah. Temuan ini menunjukkan bahwa segmentasi berbasis perilaku membaca dapat memberikan gambaran yang lebih komprehensif dibandingkan pendekatan demografis dalam memahami minat baca mahasiswa.

Kata Kunci: minat baca, *K-Means Clustering*, PCA, segmentasi mahasiswa.

**IMPLEMENTATION OF K-MEANS CLUSTERING FOR SEGMENTATION
AND PROFILING OF READING INTERESTS OF PALANGKA RAYA
UNIVERSITY STUDENTS**

Melinda Septa Natalina (213020503085)

Informatics Engineering Departement, Faculty of Engineering,

Palangka Raya University

Tunjung Nyaho Campus, Jl. Yos Sudarso Palangka Raya 73112

Email : melindaseptaa23@gmail.com

ABSTRACT

This study aims to classify Universitas Palangka Raya students based on their reading interest characteristics to support the development of more targeted literacy programs and library services. The study employed the K-Means Clustering method combined with Principal Component Analysis (PCA) using questionnaire data collected from 394 students. The research stages included data preprocessing, normalization using StandardScaler, dimensionality reduction using PCA, determination of the optimal number of clusters using the Elbow Method, and cluster evaluation using the Silhouette Score and Davies-Bouldin Index. PCA reduced 40 features into 16 principal components, while the clustering process produced four clusters: Active Readers, Selective Readers, Casual Readers, and Passive Readers. The Silhouette Score of 0.2125 and Davies-Bouldin Index of 1.8785 indicate a moderate clustering quality. The segmentation results revealed that 61.2% of students belonged to the Casual Reader and Passive Reader groups, characterized by relatively low reading activity. These findings suggest that behavior-based segmentation provides a more comprehensive understanding of students' reading interests compared to demographic-based approaches.

Keywords: *reading interest, K-Means Clustering, PCA, student segmentation.*

DAFTAR ISI

ABSTRAK.....	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	4
1.4. Tujuan.....	4
1.5. Manfaat.....	5
1.6. Sistematika Penulisan.....	5
1.7. Jadwal	7
BAB II LANDASAN TEORI.....	8
2.1. Tinjauan Pustaka	8
2.2 Minat Membaca.....	13
2.3. Perpustakaan Perguruan Tinggi.....	14
2.4. Segmentasi dan Profiling Minat Baca	15
2.4.1. Definisi Segmentasi.....	15
2.4.2. Definisi Profiling Klaster	16
2.4.3. Contoh Aplikasi	17
2.5. Data Mining	17
2.5.1. Definisi dan Konsep Dasar	17
2.6. Stratified Random Sampling	18
2.7. Cronbach's alpha	19
2.8. Algoritma K-Means Clustering	20
2.8.1. Penentuan Jumlah Cluster (K) dengan Metode Elbow.....	23
2.9. Evaluasi Hasil Clustering.....	24
2.9.1. Metode Evaluasi Internal	25

2.9.2. Visualisasi Cluster	26
2.10. Software Pengembangan.....	27
2.11. Instrumen Penelitian	28
2.11.1 Kuesioner Minat Baca	28
2.11.2 Penentuan Sampel.....	29
2.12. Streamlit.....	30
BAB III METODE PENELITIAN	33
3.1. Jenis dan Pendekatan Penelitian	33
3.2. Lokasi, Populasi, dan Sampel Penelitian	34
3.3. Teknik Pengumpulan Data.....	37
3.4. Uji Validitas dan Reliabilitas Instrumen	44
3.5. Tahapan Penelitian.....	45
3.5.1. Persiapan dan Pengumpulan Data	46
3.5.2. Preprocessing Dataset	49
3.5.3. Pelatihan Model.....	51
3.5.4. Perancangan Interface	56
BAB IV HASIL DAN PEMBAHASAN	60
4.1. Pengujian Instrumen Penelitian	60
4.1.1. Uji Validitas Instrumen.....	60
4.1.2. Uji Reliabilitas Instrumen	63
4.2. Preprocessing Data	66
4.3. Penentuan Jumlah Cluster	69
4.3.1. Reduksi Dimensi dengan PCA.....	69
4.3.2. Evaluasi Jumlah Cluster	71
4.4. Pelatihan Model K-Means.....	74
4.4.1. Perbandingan Clustering: Fitur Asli vs PCA	77
4.5. Evaluasi Cluster.....	77
4.6. Profiling dan Visualisasi Cluster	82
4.6.1. Profil Rata-rata Cluster	82
4.6.2. Visualisasi Genre Bacaan.....	85

4.6.3.	Heatmap Profil Cluster	87
4.6.4.	Visualisasi PCA 2D	90
4.6.5.	Distribusi Fakultas dan Jenis Kelamin	92
4.7.	Penamaan Cluster	96
4.8.	Hasil Implementasi Antarmuka Sistem (Streamlit)	98
4.9.	Pembahasan	103
4.8.1.	Interpretasi Karakteristik Tiap Cluster	104
4.8.2.	Implikasi Praktis Hasil Clustering dan Rekomendasi	105
BAB V		107
KESIMPULAN DAN SARAN		107
5.1.	Kesimpulan	107
5.2.	Saran	108

DAFTAR GAMBAR

Gambar 3. 1 Diagram Alur Penelitian.....	45
Gambar 4. 1 Variansi Kumulatif PCA	71
Gambar 4. 2 Elbow Method dan Silhouette Score	74
Gambar 4. 3 Silhouette Plot per Cluster.....	80
Gambar 4. 4 Proporsi Genre Bacaan per Cluster	85
Gambar 4. 5 Heatmap Profil Cluster.....	88
Gambar 4. 6 Visualisasi Cluster Menggunakan PCA 2D	90
Gambar 4. 7 Distribusi Per Fakultas	94
Gambar 4. 8 Halaman Prediksi Cluster Mahasiswa.....	99
Gambar 4. 9 Halaman Visualisasi Clustering	100
Gambar 4. 10 Halaman Profil & Penamaan Cluster	100
Gambar 4. 11 Halaman Distribusi Demografis per Cluster	101
Gambar 4. 12 Halaman Uji Validitas & Reliabilitas Instrumen	102

DAFTAR TABEL

Tabel 1. 1 Jadwal Pelaksanaan.....	7
Tabel 2. 1 Perbandingan Tinjauan Pustaka.....	8
Tabel 2. 2 Perbandingan Algoritma K-Means Clustering	21
Tabel 4. 1 Hasil Uji Validitas Instrumen	61
Tabel 4. 2 Hasil Uji Reliabilitas Instrumen.....	64
Tabel 4. 3 Ringkasan Hasil Preprocessing Data	67
Tabel 4. 4 Kolom Sumber Bacaan Hasil One-Hot Encoding.....	67
Tabel 4. 5 Hasil Reduksi Dimensi PCA.....	69
Tabel 4. 6 Hasil Evaluasi Jumlah Cluster	72
Tabel 4. 7 Jumlah Anggota pada Setiap Cluster	75
Tabel 4. 8 Hasil Evaluasi Cluster	78
Tabel 4. 9 Profil Rata-rata Cluster (Numerik dan Likert).....	82
Tabel 4. 10 Distribusi Fakultas Responden pada Sampel Keseluruhan (N = 394).....	92
Tabel 4. 11 Distribusi Jenis Kelamin per Cluster	95
Tabel 4. 12 Ringkasan Profil Cluster Berdasarkan Penamaan.....	96

DAFTAR PUSTAKA

- Akkem, Y., Kumar, S. B., & Varanasi, A. (2023). Streamlit application for advanced ensemble learning methods in crop recommendation systems – A review and implementation. *Indian Journal of Science and Technology*, 16(48), 4688–4702.
- Ambarsari, E. W., Fathudin, D., & Alfiani, G. (2024). Utilizing K-Means clustering to understanding audience interest in SEO-optimized media content. *COMFORCH (Journal of Computer and Information Research)*, 3(2), 208–214.
- Amin, A. (2025). Personalized book recommendation model using K-Means clustering. *Journal of Machine Learning Research and Applications*, 12(1), 45–58.
- Arieska, P. K., & Herdiani, N. (2018). Pemilihan teknik sampling berdasarkan perhitungan efisiensi relatif. *Jurnal Statistika*, 6(2), 166–171.
- Balan, S., et al. (2019). *Reading Habits and Reading Attitudes among University Students. International Journal of Evaluation and Research in Education*.
- Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Dani, H., Bhople, P., Waghmare, H., Munginwar, K., & Patil, A. (2022). Review on frameworks used for deployment of machine learning model. *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, 10(2), 211–215.
- Daniel, J. (2012). *Sampling essentials: Practical guidelines for making sampling choices*. SAGE Publications.

- Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1*(2), 224–227.
- Duan, L., Wang, Y., & Chen, S. (2023). A comprehensive review of spectral clustering algorithms and their applications. *Information Sciences, 590*, 112–130.
- Ezugwu, A. E., Ikotun, A. M., Oyelade, O. O., Abualigah, L., Agushaka, J. O., Eke, C. I., & Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence, 110*, 104743.
- Fakhri, D. A., Defit, S., & Sumijan. (2021). Optimalisasi pelayanan perpustakaan terhadap minat baca menggunakan metode K-Means clustering. *Jurnal Informasi dan Teknologi, 3*(3), 160–166.
- Farahnakian, F., Nicolas, F., Farahnakian, F., Nevalainen, P., Sheikh, J., Heikkonen, J., & Raduly-Baka, C. (2023). A comprehensive study of clustering-based techniques for detecting abnormal vessel behavior. *Remote Sensing, 15*(6), 1477.
- Fränti, P., & Sieranoja, S. (2019). How much can k-means be improved by using better initialization and repeats? *Pattern Recognition, 93*, 95–112.
- Gormley, I. C., Murphy, T. B., & Raftery, A. E. (2023). Model-based clustering. *Annual Review of Statistics and Its Application, 10*, 573–595.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.
- Hamdini. (2018). *Reading Habits of University Students in Pekanbaru*. Universitas Islam Riau.

- Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
- Howell, P., Smith, J., & Davis, M. (2020). Modern approaches to data clustering and dimensionality reduction. *Data Mining and Knowledge Discovery*, 34(4), 1021–1045.
- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178–210.
- Ikotun, A. M., Habyarimana, F., & Ezugwu, A. E. (2025). Cluster validity indices for automatic clustering: A comprehensive review. *Heliyon*, 11(2), e41953.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666.
- Kamus Besar Bahasa Indonesia (KBBI). (2023). *Kamus Besar Bahasa Indonesia (KBBI) Daring*. Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. <https://kbbi.kemdikbud.go.id>
- Kasa, S. R., & Rajan, V. (2023). Avoiding inferior clusterings with misspecified Gaussian mixture models. *Scientific Reports*, 13(1), 19164.
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of cluster in K-means clustering. *International Journal of Advance Research in Computer Science and Management Studies*, 1(6), 90–95.
- Kotu, V., & Deshpande, B. (2019). *Data science: Concepts and practice* (2nd ed.). Morgan Kaufmann.
- Li, X., Zhang, Y., & Wang, H. (2024). Enhanced DBSCAN algorithm for density-based clustering in complex datasets. *Pattern Recognition*, 145, 109876.

- Liu, Y., Chen, H., & Li, Z. (2023). Optimizing DBSCAN parameters using evolutionary algorithms for spatial data analysis. *Expert Systems with Applications*, 215, 119201.
- Mondal, S., Das, A., & Kumar, R. (2024). Scalable spectral clustering techniques for large-scale data processing. *IEEE Transactions on Knowledge and Data Engineering*, 36(3), 1145–1158.
- More, V., Srivastav, S., Gaud, S., Suman, S., & Soni, P. (2024). Multiple disease prediction system using Streamlit. *TIJER - International Research Journal*, 11(4), b45–b47.
- Mumpuni, A., & Nurbaeti, R. U. (2020). Dongeng antikorupsi sebagai bahan literasi bagi peserta didik di sekolah dasar. *Jurnal Ilmiah Wahana Pendidikan*, 6(4), 573–587.
- Nasution, A., Siregar, M., & Harahap, D. (2025). Segmentasi minat baca mahasiswa menggunakan algoritma K-Means clustering. *Jurnal Ilmu Komputer dan Informasi*, 18(1), 22–31.
- Patil, S., & Loksha, V. (2022). Live Twitter sentiment analysis using Streamlit framework. *International Journal of Computer Applications*, 183(43), 15–20.
- Patil, V., Sutar, S., Ghadage, S., & Palkar, S. (2023). Gesture recognition for media interaction: A Streamlit implementation with OpenCV and MediaPipe. *International Journal of Engineering Research & Technology (IJERT)*, 12(5), 412–418.
- Pribadi, R., Santoso, B., & Wijaya, A. (2025). Implementasi machine learning untuk klasterisasi data pengunjung perpustakaan. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 12(2), 150–158.

- Ran, X., Li, J., & Wang, C. (2023). A new approach to agglomerative hierarchical clustering based on dynamic distance metrics. *Knowledge-Based Systems*, 268, 110467.
- Rane, N., et al. (2025). *Reading Habits Among Students of Universitas Atma Jaya Yogyakarta. Journal of Language Teaching and Learning (LATTE)*.
- Revathy, S., Sindhuja, M., & Jayashree, K. (2024). Parkinson's disease detection application based on SVM using Streamlit. *Journal of Healthcare Informatics Research*, 8(2), 180–195.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65.
- Salminen, J., Jansen, B. J., & An, J. (2023). A review of marketing segmentation: Methods and applications in the digital age. *Journal of Business Research*, 160, 113765.
- Santoso, B. (2018). *Data mining: Teknik pemanfaatan data untuk keperluan bisnis*. Graha Ilmu.
- Sigit, W. (2024). Penerapan K-Means untuk analisis pola peminjaman buku perpustakaan. *Jurnal Sistem Informasi dan Komputer*, 11(1), 45–52.
- Siswati. (2010). Minat membaca pada mahasiswa (Studi deskriptif pada mahasiswa Fakultas Psikologi). *Jurnal Psikologi Undip*, 8(2), 124–134.
- Sutarno, N. S. (2006). *Manajemen perpustakaan: Suatu pendekatan praktik*. Sagung Seto.
- Taherdoost, H. (2016). Sampling methods in research methodology: How to choose a sampling technique for research. *International Journal of Academic Research in Management*, 5(2), 18–27.

- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55.
- Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.
- Xu, D., & Tian, Y. (2015). A comprehensive survey of clustering algorithms. *Annals of Data Science*, 2(2), 165–193.
- Zadedehbalaei, A., Bagheri, A., & Afshar, H. (2021). A study on DBSCAN Clustering algorithm issues and a survey on its improvements. *Soft Computing Journal*, 6(1), 2–37.
- Zhao, H., Yang, Y., & Chen, L. (2020). Recent advances in deep clustering algorithms. *Neural Networks*, 128, 312–325.