

---

## PENERAPAN ALGORITMA *K-MEANS CLUSTERING* UNTUK SEGMENTASI BERDASARKAN KESIAPAN SELEKSI: *SYSTEMATIC LITERATUR REVIEW*

Imam Fakhri Muhammad<sup>1</sup>, Billy Hendrik<sup>2</sup>

Universitas Putra Indonesia “YPTK” Padang, Padang

e-mail: <sup>1</sup>imamfakhrimuhammad98@gmail.com, <sup>2</sup>billy\_hendrik@upiptk.ac.id

**Abstract:** *The application of the K-means clustering algorithm as an innovative method for readiness segmentation, enabling the grouping of individuals based on their preparedness in the selection process. With the background issue of the lack of systematic methods to identify individual readiness, this study aims to develop an effective segmentation model using K-Means. The research method involves collecting relevant data and applying the K-Means algorithm to identify optimal clusters. The results show that the K-Means algorithm successfully grouped individuals into clusters according to their selection readiness, providing strategic insights for decision-makers in conducting a more accurate and efficient selection process. This study makes a significant contribution to the development of more targeted and effective selection strategies.*

**Keyword:** *Selection Strategy; Readiness-Based Grouping; Data Clustering.*

**Abstrak:** Penerapan algoritma *K-means clustering* sebagai metode inovatif untuk segmentasi kesiapan seleksi, yang memungkinkan pengelompokan individu berdasarkan kesiapan mereka dalam proses seleksi. Dengan latar belakang permasalahan kurangnya metode sistematis untuk mengidentifikasi kesiapan individu, penelitian ini bertujuan untuk mengembangkan model segmentasi yang efektif menggunakan *K-Means*. Metode penelitian melibatkan pengumpulan data relevan dan penerapan algoritma *K-Means* untuk mengidentifikasi *cluster* yang optimal. Hasil penelitian menunjukkan bahwa algoritma *K-Means* berhasil mengelompokkan individu ke dalam *cluster* yang sesuai dengan kesiapan seleksi mereka, memberikan wawasan strategis bagi pengambil keputusan dalam proses seleksi yang lebih akurat dan efisien. Penelitian ini memberikan kontribusi signifikan dalam pengembangan strategi seleksi yang lebih tepat sasaran dan efektif.

**Kata kunci:** Strategi Seleksi; Pengelompokan Berdasarkan Kesiapan; Klasterisasi Data.

### PENDAHULUAN

Seleksi ikatan dinas di Indonesia merupakan proses kompetitif yang membutuhkan persiapan matang melalui bimbingan belajar khusus. Namun, lembaga bimbel sering menghadapi tantangan dalam menyusun program yang tepat sasaran karena keragaman tingkat kesiapan peserta (Nopriandi & Aprizal, 2024). Data menunjukkan bahwa hanya 20-30% peserta yang benar-benar siap menghadapi tahapan seleksi, sementara sisanya membutuhkan pendekatan berbeda. Kondisi ini memerlukan solusi sistematis untuk mengelompokkan peserta

berdasarkan karakteristik kesiapan mereka. Algoritma *K-means clustering* telah terbukti efektif dalam berbagai studi pendidikan, termasuk segmentasi penerima beasiswa dan analisis prestasi siswa (Abdussalam Amrullah et al., 2022). Teknik ini mampu mengelompokkan data multidimensi seperti nilai akademik, kehadiran, dan hasil tryout menjadi kluster-kluster homogen. Penelitian oleh (Andi Akram Nur Risal et al., 2024) menunjukkan akurasi hingga 85% dalam pengelompokan siswa berdasarkan kemampuan menggunakan *K-Means*. Pendekatan ini belum banyak

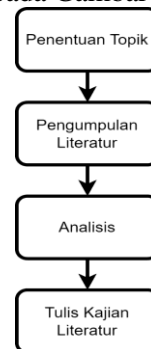
diaplikasikan khusus untuk persiapan seleksi ikatan dinas.

Berdasarkan studi pendahuluan di tiga lembaga bimbingan di Padang, ditemukan tiga masalah utama: (1) materi pembelajaran yang seragam untuk semua tingkat kemampuan, (2) ketidakefisienan alokasi tutor, dan (3) kurangnya pemetaan awal terhadap potensi peserta. Hal ini menyebabkan rendahnya persentase kelulusan peserta dengan kemampuan menengah ke bawah. Segmentasi berbasis data menjadi solusi potensial untuk mengatasi masalah ini (Hendrastuty, 2024). Penelitian ini bertujuan untuk: (1) mengimplementasikan algoritma K-Means dalam mengelompokkan peserta bimbingan berdasarkan parameter kesiapan, (2) menganalisis variabel dominan yang mempengaruhi pengelompokan, dan (3) menyusun rekomendasi program bimbingan yang disesuaikan dengan karakteristik setiap kluster. Hasil penelitian diharapkan dapat meningkatkan efektivitas program persiapan seleksi. Studi ini memberikan kontribusi ganda berupa: (1) pengembangan model *clustering* khusus untuk persiapan seleksi ikatan dinas yang belum banyak diteliti, dan (2) solusi praktis bagi lembaga bimbingan dalam menyusun program yang lebih terpersonalisasi. Secara akademis, penelitian ini memperkaya literatur tentang penerapan data mining dalam pendidikan non-formal di Indonesia (Fimawahib & Rouza, 2021)

## METODE

Studi literatur diawali dengan mengidentifikasi dan mengakses berbagai artikel penelitian melalui platform pencarian akademis. Proses ini dilakukan secara daring dengan menelusuri sumber-sumber terpercaya, terutama artikel yang dipublikasikan di Indonesia, guna memperoleh referensi yang relevan dan mendukung penelitian (Eka et al., 2022). Tahapan dalam kajian literatur ini meliputi penentuan topik penelitian, pengumpulan literatur yang relevan,

analisis literatur dengan membaca secara menyeluruh, pengelompokan berdasarkan metode dan kriteria yang digunakan dalam penelitian sebelumnya, hingga penyusunan tinjauan literatur (Casro et al., 2020). Model dari langkah-langkah kajian literatur yang dilakukan dapat dilihat pada Gambar 1 dibawah ini:



**Gambar 1. Langkah-Langkah Literatur**

Gambar 1 diatas menunjukkan prosedur dari tahapan langkah-langkah kajian literatur. Penggunaan konsep tersebut digunakan agar tahapan kajian literatur dapat dilakukan secara sistematis (Maudi Pangestu et al., 2023). Berdasarkan hal tersebut dapatlah peneliti melakukan sebuah kajian literatur dengan baik dan jelas (Eka et al., 2022). Berikut penjelasan terhadap langkah-langkah tersebut berdasarkan tahapannya sebagai berikut ini.

### Penentuan Topik

Topik penelitian ini muncul dari observasi terhadap tingginya tingkat kegagalan peserta seleksi ikatan dinas yang mempersiapkan diri secara mandiri tanpa mengikuti bimbingan, disebabkan oleh ketiadaan alat untuk mengukur kesiapan belajar secara objektif dan ketidakterediaan panduan belajar yang terpersonalisasi. Melihat efektivitas algoritma K-Means dalam berbagai studi segmentasi pendidikan, penulis memilih untuk mengaplikasikan metode ini guna mengelompokkan calon peserta berdasarkan tingkat kesiapan mandiri mereka dengan parameter seperti riwayat akademik, hasil tryout mandiri, konsistensi belajar, dan akses terhadap

sumber belajar, sehingga dapat menghasilkan rekomendasi strategi persiapan yang tepat sasaran sesuai karakteristik masing-masing kluster (Basalamah & Setyadi, 2023). Pemilihan topik ini diharapkan dapat memberikan solusi inovatif bagi masalah kesenjangan akses terhadap bimbingan belajar profesional, khususnya bagi peserta dari daerah dengan sumber daya terbatas.

### Pengumpulan Literatur

Proses ini mencakup pengumpulan artikel ilmiah yang membahas penerapan *K-means clustering* dalam konteks pendidikan dan seleksi kompetitif, menggunakan kata kunci seperti "*K-means clustering* pendidikan", "segmentasi kesiapan belajar mandiri", dan "seleksi ikatan dinas" pada *database* seperti Google Scholar, Garuda Kemdikbud, dan repositori universitas. Referensi utama meliputi studi tentang klusterisasi minat siswa dalam pemilihan sekolah, distribusi bantuan pendidikan berbasis *K-Means*, serta seleksi penerima beasiswa, yang menunjukkan konsistensi algoritma ini dalam mengelompokkan data pendidikan berdasarkan parameter multidimensi seperti prestasi akademik, akses sumber belajar, dan kebutuhan spesifik

### Analisis

Peneliti melakukan evaluasi penerapan kesiapan terhadap metode dan hasil penelitian terdahulu mengenai penerapan algoritma *K-means clustering* dalam kesiapan siswa dalam menghadapi seleksi. Analisis ini mencakup beberapa aspek utama yang berkaitan dengan efisiensi, efektivitas, dan kriteria evaluasi dalam kesiapan ini (Mauliadi, 2022). Aspek-aspek tersebut meliputi:

1. Efisiensi penerapan algoritma *K-means clustering* dalam kesiapan siswa.
2. Efektivitas implementasi diuji melalui berbagai studi kasus praktis untuk menilai kinerjanya dalam situasi nyata.

3. Kriteria evaluasi seperti variabel data hasil tryout siswa (nilai ujian akademik, psikotest), kehadiran di kelas, Tingkat pemahaman siswa berdasarkan simulasi ujian, durasi belajar siswa dalam platform digital.

### Tulis Kajian Literatur

Tahap akhir penelitian ini adalah menyusun temuan dari tinjauan pustaka secara sistematis. Peneliti menjelaskan cara kerja penerapan *K-means clustering* dalam kesiapan siswa. Serta memberikan rekomendasi strategi bimbingan bagi tiap segmen siswa. Aspek-aspek yang dicakup dalam algoritma *K-means clustering* meliputi:

1. Konteks segeentasi siswa dalam kesiapan mengikuti iktan dinas
2. Hasil implementasi dalam berbagai kasus.
3. Metodologi (algoritma *k-means clustering*).
4. Hasil implementasi diterapkan dalam berbagai studi kasus.

## HASIL DAN PEMBAHASAN

Hasil penelusuran mengidentifikasi lima belas artikel ilmiah yang disintesis secara naratif sesuai dengan kriteria inklusi yang ditetapkan. Rincian artikel-artikel tersebut disajikan dalam bentuk tabel. Hasil analisis artikel ilmiah terkait segemntasi *k-means clustering* dalam kesiapan seleksi seperti dapat dilihat pada Tabel 1 dibawah ini:

**Tabel 1. Hasil Analisis Artikel Ilmiah**

Penelitian	Hasil	Metodologi
Pemanfaatan Algoritma <i>K-means clustering</i> untuk Analisis Penilaian Prestasi Belajar Siswa di SMK	K-Means berhasil mengelompokkan siswa ke dalam 3 kluster berdasarkan prestasi akademis	<i>K-means clustering</i>

Penelitian	Hasil	Metodologi	Penelitian	Hasil	Metodologi
Negeri 1 Gowa				n jumlah rombel, jumlah peserta didik, dan jumlah tenaga guru & tendik	
Implementasi Metode Algoritma <i>K-means clustering</i> Untuk Menentukan Penerima Program Pintar (PIP)	Berhasil mengelompokkan data siswa ke dalam 3 <i>cluster</i> berdasarkan jumlah siswa per kabupaten/kota pada jenjang SD, SMP, SMA, dan Perguruan Tinggi. Hasil <i>clustering</i> : <i>Cluster</i> 0: 16 data <i>Cluster</i> 1: 4 data <i>Cluster</i> 2: 15 data	<i>K-means clustering</i>	Analisis <i>K-means clustering</i> Angkatan Kerja Berdasarkan Jenis Kelamin, Usia, Pendidikan, dan Pendapatan pada Data Satuan Kerja Angkatan Nasional Tahun 2021	- Menggunkan data SAKERN AS 2021 dengan 5 variabel utama	<i>K-means clustering</i>
Klasterisasi Penentuan Minat Siswa dalam Pemilihan Sekolah Menggunakan Metode Algoritma <i>K-means clustering</i>	- Dari 10 data sekolah SMP, diperoleh 3 <i>cluster</i> : - <i>Cluster</i> 0 (kurang diminati): 6 sekolah - <i>Cluster</i> 1 (cukup diminati): 3 sekolah - <i>Cluster</i> 2 (sangat diminati): 1 sekolah - Pengelompokan berdasarkan	- Algoritma <i>K-means clustering</i>	Analisis Suksesi SDM Menggunakan Algoritma Klasifikasi K-Nearest Neighbour dan Algoritma <i>Clustering</i> K-Means (Studi Kasus: Mabes Polri)	Algoritma K-NN menghasilkan akurasi 77.00% untuk klasifikasi kenaikan pangkat	Algoritma K-NN dan Algoritma <i>Clustering</i> K-Means
			Pemanfaatan Algoritma <i>K-means clustering</i> dalam Penentuan	Mengelompokkan data siswa ke dalam 3 <i>cluster</i> berdasarkan	Algoritma <i>K-means clustering</i>

Penelitian	Hasil	Metodologi	Penelitian	Hasil	Metodologi
Prioritas Penerima Program Bantuan Sosial Pendidikan	n jumlah siswa per kabupaten/kota pada jenjang SD, SMP, SMA, dan Perguruan Tinggi		Metode <i>K-means clustering</i> Untuk Seleksi Atlet Taekwondo Porprov	mengelompokkan atlet taekwondo ke dalam 3 <i>cluster</i> berdasarkan umur, berat badan, kuda-kuda, tendangan, dan pukulan	<i>K-means clustering</i>
Penerapan Data Mining Clustering Algoritma K-Means Untuk Menganalisa Pola Kejadian Tindak Kejahatan (Studi Kasus Polirestabes Semarang)	Algoritma <i>K-means clustering</i> berhasil mengelompokkan 305 data kejadian tindak kejahatan di Kota Semarang tahun 2022 menjadi 7 <i>cluster</i>	Algoritma <i>K-means clustering</i>	Penerapan <i>K-means clustering</i> pada Penentuan Jenis Pembelajaran di Universitas Pasir Pengaraian	Algoritma K-Means berhasil mengelompokkan 120 data dosen ke dalam 3 <i>cluster</i> : <i>Cluster 0</i> (Offline): 49 dosen, <i>Cluster 1</i> (Online): 17 dosen, <i>Cluster 2</i> (Blended Learning): 54 dosen	<i>K-means clustering</i>
Pengelompokan Bidang Keahlian Mahasiswa Menggunakan Algoritma K-Means (Studi Kasus: STMIK Kaputama Binjai)	Algoritma K-Means berhasil mengelompokkan mahasiswa ke dalam 3 kluster berdasarkan program studi, IPK, dan bidang keahlian	Algoritma K-Means	Algoritma <i>K-means clustering</i> untuk Segmentasi Pelanggan (Studi Kasus: Fashion Viral Solo)	Berhasil mengelompokkan 351 pelanggan ke dalam 2 <i>cluster</i> : <i>Cluster 1</i> memiliki nilai rata-rata Monetary Rp. 286.780, Frequency 2,34, dan Recency	Algoritma <i>K-means clustering</i>
Segmentasi Pelanggan Menggunakan <i>K-means clustering</i> di Toko Retail	Algoritma K-Means berhasil mengelompokkan data pelanggan menjadi 6 <i>cluster</i> optimal	<i>K-means clustering</i>	Penerapan	Berhasil	Metode

Penelitian	Hasil	Metodologi	Penelitian	Hasil	Metodologi
	47,69 hari. <i>Cluster 2</i> memiliki nilai rata-rata Monetary Rp. 5.002.197, Frequency 5,50, dan Recency 47,25 hari		Mining <i>Clustering</i> Menggunakan Metode K-Means Pada Data Tindak Kriminalitas Di Polres Kabupaten Kuningan	pokkan 800 data tindak kriminalitas ke dalam 9 <i>cluster</i> . <i>Cluster 0</i> : 141 items, <i>Cluster 1</i> : 128 items, <i>Cluster 2</i> : 106 items, <i>Cluster 3</i> : 110 items, <i>Cluster 5</i> : 151 items, <i>Cluster 6</i> : 49 items, <i>Cluster 7</i> : 45 items, <i>Cluster 8</i> : 2 items	<i>clustering</i>
Penerapan Algoritma <i>K-means clustering</i> dalam Pengelompokan Arsip SKCK	Berhasil mengelompokkan 416 data arsip SKCK ke dalam 3 <i>cluster</i> : <i>Cluster 1</i> : 124 data, <i>Cluster 2</i> : 210 data, <i>Cluster 3</i> : 82 data	Algoritma <i>K-means clustering</i>			
Segmentasi Siswa Menggunakan <i>K-means clustering</i> Studi Kasus Sekolah Dasar Tahfizh Quran (SDTQ) Citamulia	Berhasil mengelompokkan data siswa ke dalam 4 <i>cluster</i> : <i>Cluster 1</i> (Jarak Terjangkau): 28 siswa, <i>Cluster 2</i> (Jarak Jauh): 5 siswa, <i>Cluster 3</i> (Tetangga): 92 siswa, <i>Cluster 4</i> (Berkemampuan): 82 siswa	<i>K-means clustering</i>			
Penerapan Data	- Berhasil mengelompokkan	Metode <i>K-means</i>			

Temuan dalam studi ini menunjukkan bahwa algoritma *K-Means Clustering* sangat potensial untuk digunakan dalam mengidentifikasi segmentasi individu berdasarkan tingkat kesiapan mereka. Keunggulan utama dari metode ini adalah kemampuannya untuk memproses data multivariat dan menghasilkan pengelompokan yang jelas serta mudah diinterpretasikan.

Dari tinjauan literatur, beberapa kelebihan dan tantangan utama dalam implementasi *K-Means* untuk segmentasi kesiapan dapat dirangkum sebagai berikut:

**Kelebihan:**

1. Sederhana dan Efisien: Proses pelatihan cepat meskipun dataset besar.
2. Interpretatif: Hasil kluster dapat dijelaskan secara logis berdasarkan karakteristik kesiapan.
3. Fleksibel: Dapat digunakan pada berbagai domain seperti pendidikan, HRD, hingga pengambilan keputusan kebijakan.

**Tantangan:**

1. Sensitif terhadap outlier: K-Means dapat menghasilkan kluster yang tidak representatif jika data mengandung nilai ekstrim.
2. Penentuan jumlah kluster: Memilih jumlah  $k$  yang optimal memerlukan eksperimen dan evaluasi tambahan.
3. Skalabilitas: Untuk data yang sangat besar atau kompleks, diperlukan praproses dan normalisasi yang hati-hati.

Secara keseluruhan, literatur menunjukkan bahwa *K-Means Clustering* mampu menjadi alat bantu pengambilan keputusan yang efektif dalam konteks segmentasi kesiapan. Penggunaan metode ini mendukung strategi seleksi yang lebih adaptif, personal, dan berbasis data.

## SIMPULAN

Berdasarkan hasil kajian sistematis terhadap berbagai literatur yang relevan, dapat disimpulkan bahwa algoritma *K-Means Clustering* merupakan metode yang efektif dan efisien dalam melakukan segmentasi individu berdasarkan tingkat kesiapan mereka dalam suatu proses seleksi. K-Means mampu mengelompokkan individu ke dalam beberapa kluster dengan karakteristik kesiapan yang berbeda, sehingga dapat membantu pengambil keputusan dalam merancang strategi seleksi yang lebih tepat sasaran dan berbasis data.

Penerapan algoritma ini telah terbukti bermanfaat dalam berbagai bidang, seperti pendidikan, rekrutmen tenaga kerja, dan pelatihan sumber daya manusia. Meskipun demikian, implementasi K-Means tetap memerlukan perhatian terhadap validitas data, pemilihan jumlah kluster yang tepat, serta penanganan terhadap outlier yang dapat memengaruhi hasil klusterisasi. Secara umum, penelitian ini memberikan kontribusi penting dalam menegaskan peran *K-Means* sebagai pendekatan inovatif dalam mendukung proses seleksi yang lebih objektif, adaptif, dan efisien

## DAFTAR PUSTAKA

- Abdussalam Amrullah, Intam Purnamasari, Betha Nurina Sari, Garno, & Apriade Voutama. (2022). Analisis *Cluster* Faktor Penunjang Pendidikan Menggunakan Algoritma K-Means (Studi Kasus: Kabupaten Karawang). *Jurnal Informatika Dan Rekayasa Elektronik*, 5(2), 244–252. <https://doi.org/10.36595/jire.v5i2.701>
- Andi Akram Nur Risal, Dyah Darma Andayani, Muh Ilham Suherman, & Andi Baso Kaswar. (2024). Utilizing the K-Means *Clustering* Algorithm for Analyzing Student Achievement Assessment at SMK Negeri 1 Gowa. *Journal of Embedded Systems, Security and Intelligent Systems*, 05(March), 60–67. <https://doi.org/10.59562/jessi.v5i1.2178>
- Basalamah, A. T., & Setyadi, R. (2023). Penerapan Algoritma K-Means *Clustering* Pada Tingkat Penyelesaian Pendidikan Di Provinsi Indonesia. *Jurnal Informatika Dan Teknologi Komputer*, 4(2), 114–121. <https://ejurnalunsam.id/index.php/jicom/>
- Casro, C., Purwati, Y., Setyaningsih, G., & Kuncoro, A. P. (2020). Rancang Bangun Aplikasi Pengaduan Pelanggan Berbasis Web Menggunakan Framework Codeigniter Di Indotekno Purwokerto. *Jurnal Sains Dan Informatika*, 6(2), 166–174. <https://doi.org/10.34128/jsi.v6i2.244>
- Eka, I. P., Nugraha, I., & Lampu, W. (2022). Sistem Informasi Pemesanan Lapangan Futsal Berbasis Android Studi Kasus Jongrang Futsal. *Manajemen Dan Teknologi Informasi*, 12. <https://doi.org/10.1016/j.procs.2019.01.106%0Ahttps://doi.org/10.1016/j.apenergy.2019.114422%0Ahttp://dx.doi.org/10.1016/j.ijfatigue.2008.11.016%0Ahttp://www.ansr.pt/Estatisticas/RelatoriosDeSinistralidade/Pages/default.aspx%0Ahttp://dx.doi.org/10.1>

- 016/j.
- Fimawahib, L., & Rouza, E. (2021). Penerapan K-Means *Clustering* pada Penentuan Jenis Pembelajaran di Universitas Pasir Pengaraian. *INOVTEK Polbeng - Seri Informatika*, 6(2), 234. <https://doi.org/10.35314/isi.v6i2.2096>
- Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means *Clustering* Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (Jima-Ilkom)*, 3(1), 46–56. <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- Maudi Pangestu, K., Suendri, & Alda, M. (2023). KLIK: Kajian Ilmiah Informatika dan Komputer Sistem Informasi Geografis Sebaran Lembaga Pendidikan Qur'an (LPQ) Menggunakan Algoritma A-Star. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(2), 969–977. <https://doi.org/10.30865/klik.v4i2.1251>
- Mauliadi, R. (2022). Data Mining Menggunakan Algoritma K-Means *Clustering* dalam Analisis Tingkat Potongan Harga Terhadap Harga Jual Sepeda Motor Honda. *Jurnal Informatika Ekonomi Bisnis*, 4, 7–9. <https://doi.org/10.37034/infec.v4i4.156>
- Nopriandi, H., & Aprizal. (2024). Implementasi Sistem Pendukung Keputusan Seleksi Calon Penerima Beasiswa Kartu Indonesia Pintar Kuliah (Kip-K) Di Universitas Islam Kuantan Singingi. *ZONAsi: Jurnal Sistem Informasi*, 6(1), 123–135. <https://doi.org/10.31849/zn.v6i1.18271>