

---

## PENERAPAN ALGORITMA K-MEANS CLUSTERING PADA DATA NILAI SISWA UNTUK MENENTUKAN KELOMPOK PENERIMA BEASISWA

Sindrawati<sup>1)</sup>, Dodi Syaripudin<sup>2)</sup>, Agung Rachmat Raharja<sup>3)</sup>

<sup>1,2)</sup> Sistem Informasi, Fakultas Kesehatan dan Teknik, Universitas Bandung

<sup>3)</sup> Informatika, Fakultas Kesehatan dan Teknik, Universitas Bandung

Email: sindrawati@bandunguniversity.ac.id, dodisyaripudin@bandunguniversity.ac.id,  
agungmat@bandunguniversity.ac.id.

---

Diterima:  
26 Juli 2024

Diterima Setelah Revisi:  
7 Agustus 2024

Dipublikasikan:  
16 Agustus 2024

---

### Abstrak

Untuk mencegah terjadinya kesalahan sekolah dalam menentukan siswa kelas XII yang berhak menerima beasiswa, maka dapat dilakukan pencegahan dengan menggunakan teknik data mining sehingga pihak sekolah dapat menentukan keputusan dengan tepat dan cepat. Salah satu metode pengolahan data yang dikenal sebagai "clustering" melibatkan pengelompokan sejumlah objek atau data ke dalam kelompok, atau kelompok, sehingga setiap kelompok berisi data yang seminimal mungkin mirip dan objek yang berbeda dari kelompok lainnya. Metode yang digunakan adalah CRISP-DM, melalui proses pemahaman bisnis, pemahaman data, persiapan data, modeling, analisis, dan penerapan. K-Means adalah salah satu metode mengelompokkan data non-hierarchical yang dapat mengelompokkan data siswa ke dalam berbagai cluster berdasarkan kemiripan. Algoritma ini digunakan untuk pembentukan cluster. 109 data digunakan, dengan atribut jurusan dan nilai matematika, B. Indonesia, dan B. Inggris. Tiga cluster terbentuk; cluster pertama memiliki 26 siswa, *cluster* kedua berjumlah 46 siswa, dan *cluster* ketiga berjumlah 37 siswa. Hasil dari penelitian ini dapat digunakan sebagai salah satu dasar pengambilan keputusan untuk menentukan kelompok penerima beasiswa berdasarkan *cluster* yang terbentuk.

**Kata kunci:** Beasiswa, nilai siswa, *data mining*, *clustering*, *k-means*

### Abstract

*To prevent school mistakes in determining class XII students who are entitled to receive scholarships, prevention can be done by using data mining techniques so that the school can determine decisions accurately and quickly. Clustering is a data mining technique that functions to group a number of data or objects into clusters (groups) so that each cluster will contain data that is as similar as possible and different from objects in other clusters. The method used is CRISP-DM, through the business understanding process, data understanding, data modeling, deployment, assessment, and preparation. The K-Means algorithm is the one used to construct clusters. K-Means is a non-hierarchical technique for clustering data that divides student data into multiple clusters according to how similar the data is. A total of 109 data points were used, including majors and grades in English, Indonesian, and mathematics. The three clusters that were created are as follows: the first cluster has 26 students, the second has 46 students, and the third has 37 students. Based on the clusters that were generated, the study's findings can be utilized to inform decisions about which scholarship recipients should be grouped.*

**Keywords:** Scholarships, Student Values, Data Mining, Clustering, K-Means

## 1. PENDAHULUAN

Salah satu faktor penting dalam menghasilkan sumber daya manusia yang berkualitas tinggi adalah pendidikan. Beasiswa adalah bantuan yang diberikan oleh pihak tertentu kepada individu untuk membantu mereka terus belajar. (Jayadi, 2024). Banyak lembaga, baik pemerintah maupun non-pemerintah, telah meluncurkan program beasiswa. Calon penerima dipilih berdasarkan berbagai kriteria, termasuk nilai akademik. (Budiman & Parandani, 2019).

Muzakir (2024) menyatakan nilai siswa adalah bagian penting dari sistem pembelajaran sekolah karena nilai siswa berfungsi sebagai ukuran seberapa baik siswa menguasai materi pelajaran dan menjadi dasar pengambilan keputusan. Untuk membuat kebijakan beasiswa, sekolah dapat menggunakan hasil pengelompokan data nilai siswa untuk membedakan nilai yang baik dan buruk untuk masing-masing kelompok siswa.

Tujuan dari clustering adalah untuk mengklasifikasikan data, dengan cara menentukan pengelompokan dalam satu set data yang tidak diketahui (Wakhidah, 2023), menurut (Hendrastuty, 2024) Clustering merupakan salah satu teknik penting dalam data mining yang digunakan untuk mengelompokkan data menjadi subset yang serupa berdasarkan karakteristik atau pola tertentu. Tujuan utama dari clustering adalah untuk mengidentifikasi struktur yang tersembunyi dalam data, yang dapat membantu dalam pemahaman lebih lanjut tentang kelompok atau kategori yang ada di dalamnya.

Data dikelompokkan dengan algoritma K-Means berdasarkan centroid, atau titik pusat klaster, yang paling dekat. Tujuan algoritma ini adalah untuk meminimalkan kemiripan antar klaster dan memaksimalkan kemiripan dalam satu klaster. Dalam klaster, ukuran kemiripan berfungsi sebagai fungsi jarak. Oleh karena itu, tingkat kemiripan tertinggi dari data didapat berdasarkan jarak terkecil antara titik centroid dan data tersebut. Clustering adalah proses mengelompokkan sejumlah besar item data ke dalam grup yang lebih kecil dengan tujuan untuk memastikan bahwa masing-masing grup memiliki persamaan yang penting. (Asroni, 2023).

## 2. KAJIAN PUSTAKA

### 2.1 Algoritma

Menurut (Mulawarman, 2024) salah atau mencapai tujuan tertentu. Algoritma dapat diibaratkan sebagai resep yang memberikan langkah-langkah yang harus diikuti untuk mencapai hasil yang diinginkan. Dalam konteks komputasi, algoritma diimplementasikan melalui bahasa pemrograman yang memungkinkan komputer untuk menjalankan instruksi tersebut. Setiap algoritma memiliki karakteristik tertentu, seperti input, output, kejelasan, finiteness (berakhir), dan efektivitas. Contoh sederhana dari algoritma adalah instruksi untuk membuat secangkir teh, yang melibatkan langkah-langkah seperti mendidihkan air, menambahkan teh, dan menuangkan air mendidih ke dalam cangkir. Selain itu menurut (Raharja et al., 2024) Algoritma adalah rangkaian langkah-langkah logis yang disusun secara sistematis dan logis untuk menyelesaikan suatu masalah."

### 2.2 K-Means Clustering

Clustering merupakan salah satu teknik dalam data mining yang bertujuan untuk mengelompokkan data ke dalam kelompok-kelompok yang memiliki kesamaan berdasarkan pada karakteristik tertentu. Teknik ini mencari struktur dalam data tanpa memerlukan label kelas yang telah ditentukan sebelumnya (Hendrastuty, 2024).

## 3. METODE PENELITIAN

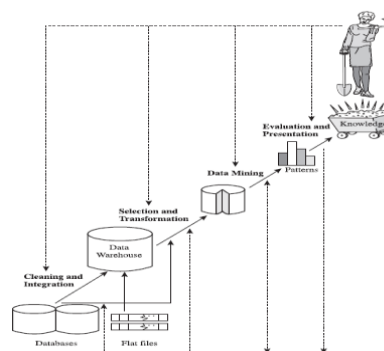
Metodologi penelitian adalah kerangka atau pendekatan sistematis yang digunakan peneliti untuk merencanakan, melakukan, dan menganalisis penelitian. Tujuan metodologi penelitian adalah untuk memastikan pelaksanaan penelitian yang sistematis, andal, dan kompeten. (Sutisna et al., 2024)

Data mining adalah proses menganalisa Data mining adalah proses untuk menemukan pola atau korelasi dari ratusan atau ribuan field dalam database yang besar dan menggabungkannya menjadi informasi penting. (Mabrur & Riani, 2012). Menurut Budi Santosa Data mining, yang juga disebut sebagai penemuan pengetahuan dalam database (KDD), adalah proses mengumpulkan dan menggunakan data historis untuk menemukan hubungan, pola, atau keteraturan dalam kumpulan data yang sangat besar. (Mahena et al, 2015).

Untuk menemukan pola baru dan bermanfaat, teknik data mining digunakan untuk memeriksa basis data yang sangat besar. Data mining bukanlah istilah untuk semua jenis pencarian informasi. Sebagai contoh, pencarian record individual melalui sistem manajemen database atau pencarian halaman web tertentu melalui queri ke semua search engine adalah tugas pencarian informasi yang erat kaitannya dengan pencarian informasi.

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat di lakukan, yaitu (Afriansyah, 2015):

1. Deskripsi  
Para peneliti biasanya mencoba menemukan cara untuk mendeskripsikan pola dan trend yang tersembunyi dalam data.
2. Estimasi  
Estimasi hampir mirip dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori.
3. Prediksi  
Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dan hasil akan ada di masa mendatang.
4. Klasifikasi  
Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori. Yaitu pendapatan tinggi, sedang, dan rendah.
5. Clustering  
Clustering menggabungkan daftar, pengamatan, atau kasus dalam kelas yang memiliki karakteristik yang sama.
6. Asosiasi  
Identifikasi hubungan antara berbagai peristiwa yang terjadi pada satu waktu.



Gambar 1. Tahapan *data mining*.  
Sumber: Han & Kamber, (2006)

Tahap-tahap data mining adalah sebagai berikut:

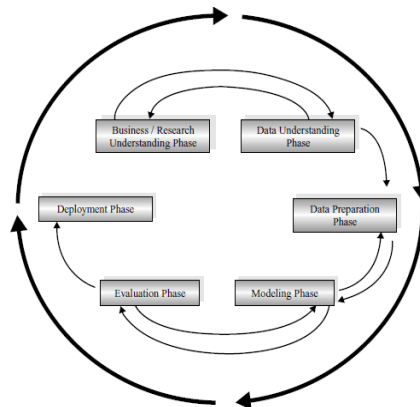
1. Pembersihan data (data cleaning)  
Pembersihan data menghapus data yang tidak konsisten dan tidak relevan..
2. Integrasi data (data integration)

Penggabungan data dari berbagai database ke dalam satu database baru dikenal sebagai integrasi data.

3. Seleksi data (data selection)  
Data database seringkali tidak digunakan sepenuhnya; oleh karena itu, hanya data yang tepat untuk dianalisis yang akan diambil dari database.
4. Transformasi data (data transformation)  
Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining.
5. Proses mining  
Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.
6. Evaluasi pola (pattern evaluation)  
Untuk mengidentifikasi pola-pola menarik ke dalam *knowledgebased* yang ditemukan.
7. Presentasi pengetahuan (knowledge presentation)  
Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

### 3.1 Metode Penelitian Data Mining

Metode Penelitian data mining dan model melanjutkan cakupan data mining sebagai suatu proses, standar proses yang digunakan adalah kerangka kerja CRISP-DM (Cross-Industri Standar process for Data Mining). Dalam CRISP-DM data mining dilihat sebagai proses keseluruhan dari komunikasi masalah bisnis hingga penerapan model (Larose, 2005).



Gambar 2. Tahapan proses CRISP-DM.  
Sumber: Larose, (2005)

Enam fase CRISP-DM (Larose, 2005) :

1. Fase Pemahaman bisnis (Business Understanding Phase)
2. Fase Pemahaman Data (Data Understanding Phase)
3. Fase Pengolahan Data (Data Preparation Phase)
4. Fase Pemodelan (Modeling Phase)
5. Fase Evaluasi (Evaluation Phase)
6. Fase Penyebaran (Deployment Phase)

### 3.2 Algoritma K-Means

K-Means adalah konstanta jumlah cluster yang diinginkan, dan Means adalah nilai rata-rata dari grup data yang disebut cluster. Dengan demikian, K-Means adalah metode penganalisaan data atau

pengolahan data yang melakukan proses pemodelan tanpa supervisi, dan juga memungkinkan pengelompokan data dengan sistem partisi. (Nasari & Darma, 2015).

Dalam metode pengelompokan, algoritma yang dikenal sebagai fungsi jarak—atau tahap persamaan atau perbedaan—digunakan untuk menghitung perbedaan kelompok. Untuk itu digunakan algoritma k-means yang di dalamnya memuat aturan sebagai berikut:

1. Jumlah cluster yang perlu dimasukkan
2. Hanya memiliki atribut dengan tipe numeric

Dasar algoritma k-means adalah sebagai berikut:

1. Tentukan K sebagai jumlah cluster yang ingin dibentuk
2. Tentukan K centroid (titik pusat cluster) awal secara random
3. Hitung jarak setiap data ke masing-masing centroids menggunakan rumus Euclidean Distance
4. Kelompok setiap data berdasarkan jarak terdekat antara data dengan centroidnya
5. Tentukan posisi centroids baru (k C) dengan cara menghitung nilai rata-rata dari data yang terletak pada centroids yang sama.

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

6. Kembali ke langkah 3, jika posisi centroid baru dengan centroid lama tidak sama.
7. Iterasi berhenti dan dianggap stabil jika hasil iterasi baru sama dengan hasil iterasi sebelumnya.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Pengolahan Data Awal

Pada tahap ini merupakan tahap untuk memastikan data siswa yang dipilih telah layak untuk dilakukan proses pengolahan.

#### A. Data Cleaning

Pada titik ini, atribut data yang akan digunakan dalam penelitian dipilih, dan atribut data yang tidak digunakan dibuang. seperti nama orang tua, jenis kelamin, dan tempat tanggal lahir akan dihilangkan.

#### B. Data Transformation

Data yang berjenis nominal seperti daerah asal dan jurusan harus dilakukan proses inialisasi data terlebih dahulu ke dalam bentuk angka atau numerikal. Untuk melakukan inialisasi daerah asal dapat dilakukan dengan cara pengurutan angka berdasarkan frekuensinya.

Tabel 1. Tabel Inialisasi Jurusan

Jurusan	Frekuensi	Inialisasi
IPA	75	1
IPS	34	2

Berdasarkan tabel diatas penentuan urutan number inialisasi ditentukan berdasarkan besar frekuensinya. Yaitu diurutkan berdasarkan frekuensi terbesar ke frekuensi terkecil.

Tabel 2. Contoh tabel dataset yang sudah dilakukan inialisasi

NO	JURUSAN	NILAI		
		MATEMATIKA	INDONESIA	INGGRIS
1	1	77.00	80.00	78.00
2	1	82.00	80.00	78.00
3	1	77.00	80.00	78.00
4	1	80.50	87.50	78.00
5	1	80.00	80.00	80.00
6	2	77	80	76
7	2	78	80	78
8	2	79	80	80
9	2	78	80	75
10	2	76	80	80

Hitung jarak antara objek dan centroid cluster dengan menggunakan rumus jarak geometris. Persamaan yang digunakan antara lain:

Tabel 3. Hasil perhitungan jarak objek dan centroid

No	Jurusan	matematika	indonesia	inggris	Jarak			Cluster
					C 1	C 2	C 3	
1	1	77.00	80.00	78.00	10.89725	3.741657	3	3
2	1	82.00	80.00	78.00	9.937303	7	3.741657	3
3	1	77.00	80.00	78.00	10.89725	3.741657	3	3
4	1	80.50	87.50	78.00	6.576473	10.93161	7.968689	1
5	1	80.00	80.00	80.00	8.87412	6.708204	1.414214	3
6	1	76.00	80.00	78.00	11.34681	3.605551	3.741657	2
7	1	76.00	80.00	78.00	11.34681	3.605551	3.741657	2
8	1	76.00	78.00	75.00	14.51723	0	6.244998	2
9	1	79.50	80.00	82.00	8.154753	8.077747	2.291288	3
10	1	82.00	80.00	82.00	7.92149	9.433981	3.741657	3

adalah:

$$(p, q) = \sqrt{(p1 - q1)^2 + (p2 - q2)^2 + (p3 - q3)^2}$$

Karena atribut yang digunakan berjumlah 3, persamaan ini digunakan untuk menghitung jarak dari data siswa pertama ke pusat cluster pertama.

$$\begin{aligned} (1,1) &= \sqrt{(1 - 1)^2 + (77.00 - 81.50)^2 + (80.00 - 87.50)^2 + (78.00 - 84.50)^2} \\ &= \sqrt{(0)^2 + (-4.5)^2 + (-7.5)^2 + (-6.5)^2} \\ &= \sqrt{0 + 20.25 + 56.25 + 42.25} \\ &= \sqrt{118.75} \\ &= 10.8972 \end{aligned}$$

Dari hasil perhitungan di atas di dapatkan hasil bahwa jarak data siswa pertama dengan cluster pertama adalah 10.8972

Jarak dari data siswa pertama ke pusat cluster kedua dengan persamaan:

$$\begin{aligned}(1,2) &= \sqrt{(1-1)^2 + (77.00 - 76.00)^2 + (80.00 - 78.00)^2 + (78.00 - 75.00)^2} \\ &= \sqrt{(0)^2 + (1)^2 + (2)^2 + (3)^2} \\ &= \sqrt{0 + 1 + 4 + 9} \\ &= \sqrt{14} \\ &= 3.74166\end{aligned}$$

Dari hasil perhitungan di atas di dapatkan hasil bahwa jarak data siswa pertama dengan cluster pertama adalah 3.74166.

Jarak dari data siswa pertama ke pusat cluster ketiga dengan persamaan:

$$\begin{aligned}(1,3) &= \sqrt{(1-2)^2 + (77.00 - 79)^2 + (80.00 - 80)^2 + (78.00 - 80)^2} \\ &= \sqrt{(-1)^2 + (-2)^2 + (0)^2 + (-2)^2} \\ &= \sqrt{1 + 4 + 0 + 4} \\ &= \sqrt{9} \\ &= 3\end{aligned}$$

Hasil perhitungan di atas menunjukkan bahwa jarak data siswa pertama dengan cluster pertama adalah 3, dan berdasarkan hasil ketiga perhitungan di atas, dapat disimpulkan bahwa siswa pertama termasuk dalam cluster 2.

Kelompokkan data ke dalam cluster dengan jarak terdekat (minimal).

Tabel 4. Tabel dengan centroid awal

No	Cluster 1	Cluster 2	Cluster 3
1	0	0	1
2	0	0	1
3	0	0	1
4	1	0	0
5	0	0	1
6	0	1	0
7	0	1	0
8	0	1	0
9	0	0	1
10	0	0	1

Setelah semua data dimasukkan ke dalam cluster yang lebih dekat, hitung pusat cluster baru berdasarkan jumlah anggota rata-rata di dalamnya.

$$C_i = \left( \frac{\sum \text{jumlah jurusan anggota } C_i}{\text{jumlah data anggota } C_i} ; \frac{\sum \text{jumlah nilai matematika } C_i}{\text{jumlah data anggota } C_i} ; \frac{\sum \text{jumlah nilai B.indonesia } C_i}{\text{jumlah data anggota } C_i} ; \frac{\sum \text{jumlah nilai B.inggris } C_i}{\text{jumlah data anggota } C_i} \right)$$

Contoh perhitungan pada cluster 1 adalah sebagai berikut:

$$\frac{1}{1}, \frac{80.50}{1}, \frac{87.50}{1}, \frac{78.00}{1}$$

Jadi, hasil centroid baru cluster 1 adalah (1, 80.50, 87.50, 78.00).

Contoh perhitungan pada cluster 2 adalah sebagai berikut:

$$\frac{3}{3}, \frac{228}{3}, \frac{238}{3}, \frac{231}{3}$$

Jadi, hasil centroid baru cluster 2 adalah (1, 76, 79.3, 77).

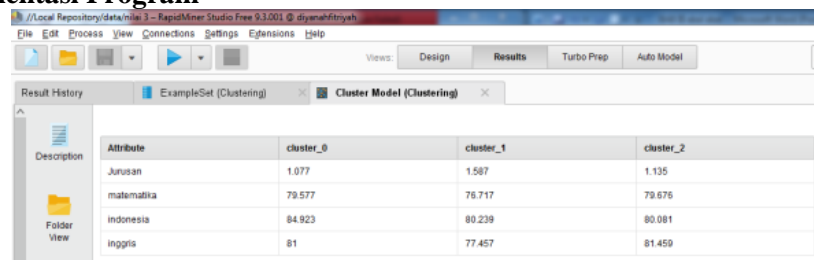
Contoh perhitungan pada cluster 3 adalah sebagai berikut:

$$\frac{6}{6}, \frac{477.5}{6}, \frac{480}{6}, \frac{478}{6}$$

Jadi, hasil centroid baru cluster 3 adalah (1, 79.6, 80, 79.6).

Setelah menciptakan titik pusat baru untuk setiap cluster, kembali ke langkah sebelumnya sampai titik pusat dari setiap cluster tidak berubah lagi dan data tidak berpindah dari satu cluster ke cluster yang lain.

#### 4.2 Implementasi Program



Attribute	cluster_0	cluster_1	cluster_2
Jurusan	1.077	1.587	1.135
matematika	79.577	76.717	79.676
Indonesia	84.923	80.239	80.081
Inggris	81	77.457	81.459

Gambar 3. Hasil rapidminer

Hasil pengujian *dataset* yang berjumlah 109 data menggunakan *software rapidminer 9.0* terbentuk 3 cluster. Pada *cluster 0* (*Cluster* pertama) terdapat 26 *items*, pada *cluster 1* (*cluster* kedua) terdapat 46 *items*, dan pada *cluster 2* (*cluster* ketiga) terdapat 37 *items*. Dari 3 *cluster* yang terbentuk kemudian di hasilkan centroid akhir untuk masing masing cluster yaitu cluster 0 = 1.077, 79.577, 84.923, 81 cluster 1 = 1.587, 76.717, 80.239, 77.457 dan cluster 2 = 1.135, 79.676, 80.081, 81.459.

Dari 3 cluster yang terbentuk, cluster 1 adalah cluster yang paling potensial untuk mendapatkan beasiswa dengan jumlah 26 siswa karena rata-rata nilainya yang tinggi. Cluster 3 dengan jumlah 37 siswa juga berpotensi mendapatkan beasiswa untuk 50% siswa dengan nilai rata-rata tertinggi. Dan di cluster 2 dengan jumlah 46 siswa tidak berpotensi untuk mendapatkan beasiswa karena nilai rata-rata yang rendah.

Hasil dari penelitian ini dapat membantu pihak Sekolah menentukan siapa yang layak mendapatkan beasiswa lanjutan.

#### 5. SIMPULAN

Untuk mencegah kesalahan pihak sekolah dalam menentukan siswa kelas XII yang berhak menerima beasiswa maka dapat dilakukan pencegahan dengan menggunakan teknik data mining sehingga pihak sekolah dapat menentukan keputusan secara tepat dan cepat. Berdasarkan hasil yang didapatkan dalam penelitian ini, maka dapat diambil beberapa kesimpulan sebagai berikut:



1. Pengolahan data yang dilakukan menghasilkan kelompok penerima beasiswa berdasarkan nilai, sehingga pihak sekolah bisa lebih mudah dalam menentukan penerima beasiswa dengan teknik data mining.
2. Berdasarkan hasil pengolahan data dapat diketahui bahwa cluster 1 adalah kelompok siswa yang berhak menerima beasiswa dengan nilai rata-rata tertinggi, cluster 2 yaitu kelompok siswa yang tidak berhak menerima beasiswa karena nilai rata-rata terendah, dan cluster 3 yaitu kelompok penerima beasiswa sebesar 50% untuk siswa dengan nilai rata-rata tertinggi.

## DAFTAR PUSTAKA

- Bodnar, G. H., & Hopwood, W. S. (2006). Sistem informasi akuntansi. *Jakarta: Salemba Empat*.
- Budiman, B. (2020). PERANCANGAN APLIKASI SISTEM TINDAK LANJUT PELANGGAN PADA PT. XYZ. *SisInfo–Jurnal Sistem Informasi dan Informatika*, 2(1), 1-7.
- Triandini, E., Jayanatha, S., Indrawan, A., Putra, G. W., & Iswara, B. (2019). Metode systematic literature review untuk identifikasi platform dan metode pengembangan sistem informasi di Indonesia. *Indonesian Journal of Information Systems*, 1(2), 63-77.
- 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>
- Larose, T. D. (2005). *An Introduction to Data Mining: Vol.* <http://www>.
- Nasari, F., & Darma, S. (2015). Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru. *Seminar Nasional Teknologi Informasi dan Multimedia 2015*, 73–78.
- Sutisna, T., Raharja, A. R., Hariyadi, E., Hafizh, V., & Putra, C. (2024). Penggunaan Computer Vision untuk Menghitung Jumlah Kendaraan dengan Menggunakan Metode SSD ( Single Shoot Detector ). *Journal Of Social Science Research Volume*, 4, 6060–6067. <https://doi.org/10.31004/innovative.v4i2.10071>
- Wakhidah, N. (2023). Clustering Menggunakan K-Means Algorithm. *Jurnal Transformatika*, 8(1), 33. <https://doi.org/10.26623/transformatika.v8i1.45>
- Jayadi, J., Raharja, A. R., Pramudianto, A., & Muchsam, Y. (2024). Application of Naïve Bayes Classifier Algorithm for Classification of Scholarship Recipients at SMA PGRI 2 Bandung. *International Journal of Mechanical Computational and Manufacturing Research*, 13(2), 33-41.
- Raharja, A. R. (2024). KEAMANAN JARINGAN. PENERBIT KBM INDONESIA.
- Raharja, A. R., Pramudianto, A., & Muchsam, Y. (2024). Penerapan Algoritma Decision Tree dalam Klasifikasi Data “Framingham” Untuk Menunjukkan Risiko Seseorang Terkena Penyakit Jantung dalam 10 Tahun Mendatang. *Technologia Journal*, 1(1).
- 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>
- Larose, T. D. (2005). *An Introduction to Data Mining: Vol.* <http://www>.
- Nasari, F., & Darma, S. (2015). Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru. *Seminar Nasional Teknologi Informasi dan Multimedia 2015*, 73–78.
- Raharja, A. R., Ramalinda, D., & Hariyanti, I. (2024). *ALGORITMA DAN PEMROGRAMAN MENGGUNAKAN PYTHON DENGAN APLIKASI GOOGLE COLLABS*. Mafy Media Literasi.
- Sutisna, T., Raharja, A. R., Hariyadi, E., Hafizh, V., & Putra, C. (2024). Penggunaan Computer Vision untuk Menghitung Jumlah Kendaraan dengan Menggunakan Metode SSD ( Single Shoot Detector ). *Journal Of Social Science Research Volume*, 4, 6060–6067. <https://doi.org/10.31004/innovative.v4i2.10071>