

Optimasi Pengelompokan Tindak Kejahatan di Provinsi Jawa Barat melalui K-Means Clustering

by - -

Submission date: 12-Feb-2024 06:21AM (UTC+0000)

Submission ID: 224199974

File name: Kejahatan_di_Provinsi_Jawa_Barat_melalui_K-Means_Clustering.docx (722.83K)

Word count: 1529

Character count: 9749

Optimasi Pengelompokan Tindak Kejahatan di Provinsi Jawa Barat melalui K-Means Clustering

<http://dx.doi.org/10.28932/jutisi.vXiX.X>

Riwayat Artikel

Received: xx Bulan 20xx | Final Revision: xx Bulan 20xx | Accepted: xx Bulan 20xx

Creative Commons License 4.0 (CC BY – NC)



Sahl¹, Najwa Khansa'in Chasna², Marsani Asfi., M.Si.³

Program Studi Teknik Informatika, Universitas Catur Insan Cendekia

Jl. Kesambi No.202, Drajat, Kec. Kesambi, Kota Cirebon, Jawa Barat 4513, Indonesia

sahl.ti.21@cic.ac.id

najwa.chasna.ti.21@cic.ac.id

²Program studi Penulis 2, Nama Universitas Penulis 2

Alamat, Kota, Kode-pos, Negara

²email.penulis2@domain.ekstensi

¹Corresponding author: sahl.ti.21@cic.ac.id

Abstrak — Pengelompokan tindak kejahatan menjadi suatu tantangan penting dalam upaya meningkatkan keamanan dan keamanan masyarakat. Penelitian ini bertujuan untuk mengoptimalkan pengelompokan tindak kejahatan di Provinsi Jawa Barat melalui metode K-Means Clustering. Data yang digunakan dalam penelitian ini adalah data tindak kejahatan yang terjadi dalam beberapa tahun terakhir di Provinsi Jawa Barat. Metode K-Means Clustering digunakan untuk mengelompokkan pola tindak kejahatan berdasarkan karakteristiknya.

Hasil penelitian menunjukkan bahwa metode K-Means Clustering mampu menghasilkan pengelompokan tindak kejahatan yang optimal berdasarkan sejumlah kriteria yang telah ditetapkan. Pengelompokan ini dapat memberikan wawasan yang berharga bagi pihak berwenang dalam merencanakan strategi penanggulangan kejahatan di Provinsi Jawa Barat. Selain itu, penelitian ini juga mengidentifikasi pola tindak kejahatan yang dominan dalam setiap kelompok, sehingga memungkinkan untuk pengembangan strategi penanganan yang lebih efektif.

Kata kunci— Clustering; K-Means; Tindak Kejahatan

Optimization of Crime Grouping in West Java Province through K-Means Clustering

JuTISI

Jurnal Teknik Informatika dan Sistem Informasi

Abstract — The grouping of crimes is an important challenge in efforts to improve public security and security. This study aims to optimize the grouping of crimes in West Java Province through the K-Means Clustering method. The data used in this study is data on crimes that occurred in recent years in West Java Province. The K-Means Clustering method is used to group crime patterns based on their characteristics.

The results showed that the K-Means Clustering method was able to produce optimal crime grouping based on a number of predetermined criteria. This grouping can provide valuable insights for authorities in planning crime reduction strategies in West Java Province.

In addition, the study also identified the dominant pattern of crime in each group, allowing for the development of more effective coping strategies.

Keywords — Clustering; K-Means; Crime

I. PENDAHULUAN

Tindak kejahatan merupakan fenomena kompleks yang memerlukan pemahaman mendalam untuk dapat mengembangkan strategi penanganan yang efektif. Provinsi Jawa Barat, sebagai salah satu wilayah terpadat di Indonesia, menghadapi tantangan serius terkait peningkatan jumlah tindak kejahatan. Oleh karena itu, perlu adanya pendekatan yang canggih dan efisien untuk menganalisis pola kejahatan guna mendukung upaya penegakan hukum dan pencegahan tindak kejahatan.

Dengan banyaknya kasus kejahatan pencurian di Provinsi Jawa Barat, maka diperlukan sistem yang terkomputerisasi untuk mengelompokkan wilayah terjadinya kejahatan.

Salah satu metode yang telah terbukti efektif dalam menganalisis pola data multidimensional seperti data tindak kejahatan adalah K-Means Clustering. K-Means Clustering adalah teknik pengelompokan data yang membagi himpunan data menjadi kelompok-kelompok homogen berdasarkan kemiripan karakteristik tertentu. Dalam konteks penelitian ini, K-Means Clustering digunakan untuk mengoptimalkan pengelompokan tindak kejahatan di Provinsi Jawa Barat, sehingga dapat membantu aparat penegak hukum dan pihak terkait untuk mengambil keputusan yang lebih cerdas dan tepat sasaran.

II. METODE PENELITIAN

a. Data Mining

Data mining digunakan secara luas di berbagai bidang, termasuk industri, keuangan, cuaca, ilmu pengetahuan, dan teknologi. Contohnya, dalam industri, data mining dapat digunakan untuk menganalisis tren produksi atau pola permintaan pelanggan. Di bidang keuangan, data mining dapat membantu dalam mendeteksi pola penipuan atau mengidentifikasi peluang investasi. Dalam data mining juga terdapat metode – metode yang dapat digunakan seperti klasifikasi, clustering, regresi, seleksi variabel, dan market basket analisis.

Hal penting yang terkait dengan data mining adalah:

1. Data mining merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan data mining adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat

b. Clustering

Tujuan utama dari data mining clustering adalah untuk mengelompokkan data ke dalam kelompok-kelompok yang memiliki kesamaan internal dan perbedaan eksternal. Terdapat beberapa metode yang digunakan dalam data mining clustering, seperti K-Means, Hierarchical Clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), dan masih banyak lagi. Setiap metode memiliki kelebihan dan kelemahan masing-masing, serta cocok untuk jenis data dan kasus penggunaan tertentu.

c. **K-Means**

K-Means merupakan Algoritma clustering yang pertama kali diperkenalkan oleh James B MacQueen pada tahun 1976. Metode ini merupakan suatu metode clustering non-heirachial yang umum digunakan yang relatif sederhana untuk mengelompokkan data dalam jumlah besar

K-Means digunakan untuk mengelompokkan data menjadi beberapa kelompok berdasarkan kesamaan fitur atau atributnya. Tujuan utama dari algoritma K-Means adalah untuk membagi himpunan data menjadi k kelompok, di mana setiap data masuk ke dalam kelompok yang memiliki rerata (centroid) terdekat dengan data tersebut. Cara kerja algoritma K-Means sebagai berikut :

1. Menentukan k (nilainya bebas) sebagai jumlah cluster yang ingin dibentuk.
2. Membangkitkan nilai random untuk pusat cluster awal (centroid) sebanyak k.
3. Menghitung jarak setiap data input terhadap masing-masing centroid menggunakan rumus jarak Euclidean (Euclidean Distance) hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Persamaan Euclidian Distance:

$$d(xi, \mu_j) = \sqrt{\sum (xi - \mu_j)^2}$$

4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil).
5. Memperbaharui nilai centroid baru di peroleh dari rata-rata cluster yang bersangkutan dengan menggunakan rumus:

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j$$

6. Melakukan perulangan dari langkah 3 hingga 5, sampai anggota tiap cluster tidak ada yang berubah.

III. HASIL DAN PEMBAHASAN

a. Data Selection

Dataset yang dipakai diperoleh dari website, data tindak kejahatan dari tahun 2019-2021, berdasarkan dataset yang diambil pada website <https://opendata.jabarpov.go.id/id/dataset/daftar-status-tindak-kejahatan-berdasarkan-jenis-kriminalitas-di-jawa-barat>

b. Data View

Untuk melihat data yang diproses, dapat diakses melalui google collab berikut : https://drive.google.com/file/d/1PXvb-g_xjIYRLhmT_pOM8Z9kw6oVSfVC/view?usp=sharing

c. Processing

Dalam penelitian ini terdapat 19 data kabupaten/ kota di provinsi Jawa Barat, yang akan diproses kedalam perhitungan k-means clustering..

```
# Assuming that 'df' is your DataFrame
# Convert 'ADA' and 'tidak' to numeric values
df.replace({'ADA': 1, 'TIDAK ADA': 0}, inplace=True)
```

Gambar 1.convert values

Langkah-1 : mengkonfersikan value ADA dan TIDAK dengan numeric 0 dan 1

```
# Create a new column 'rasio_kejahatan' that is the sum of all crime columns
crime_columns = ['terdapat_penganiayaan', 'terdapat_pencurian', 'terdapat_penipuan', 'terdapat_pembakaran', 'terdapat_per
df['rasio_kejahatan'] = df[crime_columns].sum(axis=1)

# Group by 'kecamatan_kota' and sum 'rasio_kejahatan'
df_grouped = df.groupby('bps_nama_kabupaten_kota')['rasio_kejahatan'].sum().reset_index()
```

Gambar 2. Create a new columns

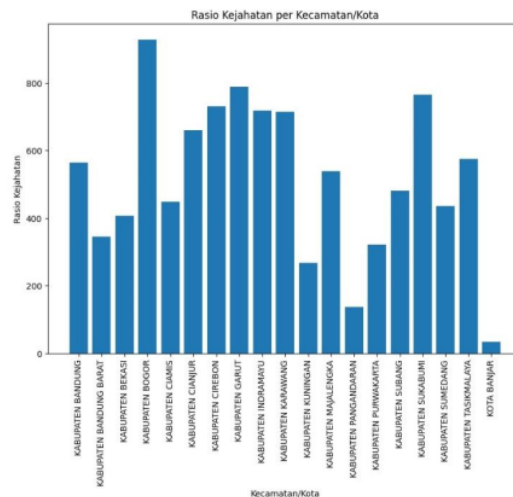
Langkah-2 : Membuat kolom 'rasio_kejahatan' yang merupakan jumlah dari semua kolom tindak kejahatan

Langkah-3 : Mengelompokkan data berdasarkan kecamatan/kota, dan menjumlahkan rasio tindak kejahatannya

```
# Plot
plt.figure(figsize=(10, 7))
plt.bar(df_grouped['bps_nama_kabupaten_kota'], df_grouped['rasio_kejahatan'])
plt.xlabel('Kecamatan/kota')
plt.ylabel('Rasio Kejahatan')
plt.title('Rasio Kejahatan per Kecamatan/Kota')
plt.xticks(rotation=90) # Rotate x-axis labels for better readability
plt.show()
```

Gambar.3 Plot

Langkah-4 : Menampilkan visual



Gambar 4. Hasil visualisasi

```
# Ubah DataFrame menjadi array numpy untuk komputasi lebih cepat
X = df_grouped['rasio_kejahatan'].values.reshape(-1, 1)

# Gunakan metode elbow untuk menemukan jumlah kluster yang optimal
wcss = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, init='k-means++', max_iter=300, n_init=10, random_state=0)
    kmeans.fit(X)
    wcss.append(kmeans.inertia_)
plt.plot(range(1, 11), wcss)
plt.title('Metode Elbow')
plt.xlabel('Jumlah kluster')
plt.ylabel('WCSS')
plt.show()
```

Gambar 5.

Langkah-4 : Mengubah data frame menjadi array numpy

Langkah-5 : Menggunakan metode elbow untuk menemukan jumlah cluster

```
# Dari plot di atas, pilih jumlah kluster yang optimal dan jalankan KMeans
n_clusters = 2 # Ganti dengan jumlah kluster yang optimal berdasarkan plot elbow
kmeans = KMeans(n_clusters=n_clusters, init='k-means++', max_iter=300, n_init=10, random_state=0)
pred_y = kmeans.fit_predict(X)

# Tambahkan label kluster ke DataFrame
df_grouped['cluster'] = pred_y

# Ganti label numerik kluster dengan label interpretatif
df_grouped['cluster'].replace({0: 'Kejahatan Rendah', 1: 'Kejahatan Tinggi'}, inplace=True)

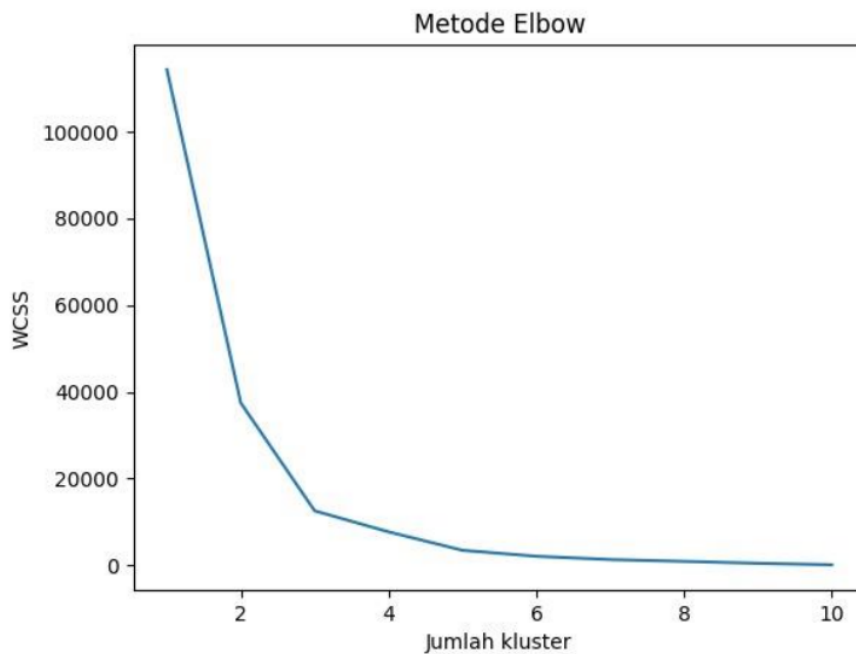
# Tampilkan hasil klustering
print(df_grouped)
```

Gambar 6.

Langkah 6-: Memilih jumlah cluster yang optimal, kemudian mengganti jumlah cluster yang optimal berdasarkan dengan plot elbow

Langkah-7 : Menambahkan label cluster ke Dataframe

Langkah-8: Mengganti label numeric cluster dengan label interpretative



Gambar 7. Hasil

	bps_nama_kabupaten_kota	rasio_kejahatan	cluster
0	KABUPATEN BANDUNG	180.0	Kejahatan Tinggi
1	KABUPATEN BANDUNG BARAT	116.0	Kejahatan Rendah
2	KABUPATEN BEKASI	134.0	Kejahatan Rendah
3	KABUPATEN BOGOR	317.0	Kejahatan Tinggi
4	KABUPATEN CIAMIS	155.0	Kejahatan Rendah
5	KABUPATEN CIANJUR	237.0	Kejahatan Tinggi
6	KABUPATEN CIREBON	237.0	Kejahatan Tinggi
7	KABUPATEN GARUT	259.0	Kejahatan Tinggi
8	KABUPATEN INDRAMAYU	235.0	Kejahatan Tinggi
9	KABUPATEN KARAWANG	234.0	Kejahatan Tinggi
10	KABUPATEN KUNINGAN	71.0	Kejahatan Rendah
11	KABUPATEN MAJALENGKA	165.0	Kejahatan Rendah
12	KABUPATEN PANGANDARAN	40.0	Kejahatan Rendah
13	KABUPATEN PURWAKARTA	116.0	Kejahatan Rendah
14	KABUPATEN SUBANG	157.0	Kejahatan Rendah
15	KABUPATEN SUKABUMI	249.0	Kejahatan Tinggi
16	KABUPATEN SUMEDANG	135.0	Kejahatan Rendah
17	KABUPATEN TASIKMALAYA	188.0	Kejahatan Tinggi
18	KOTA BANJAR	11.0	Kejahatan Rendah

Gambar 8.

Langkah-9 : Menampilkan rasio kejahatan

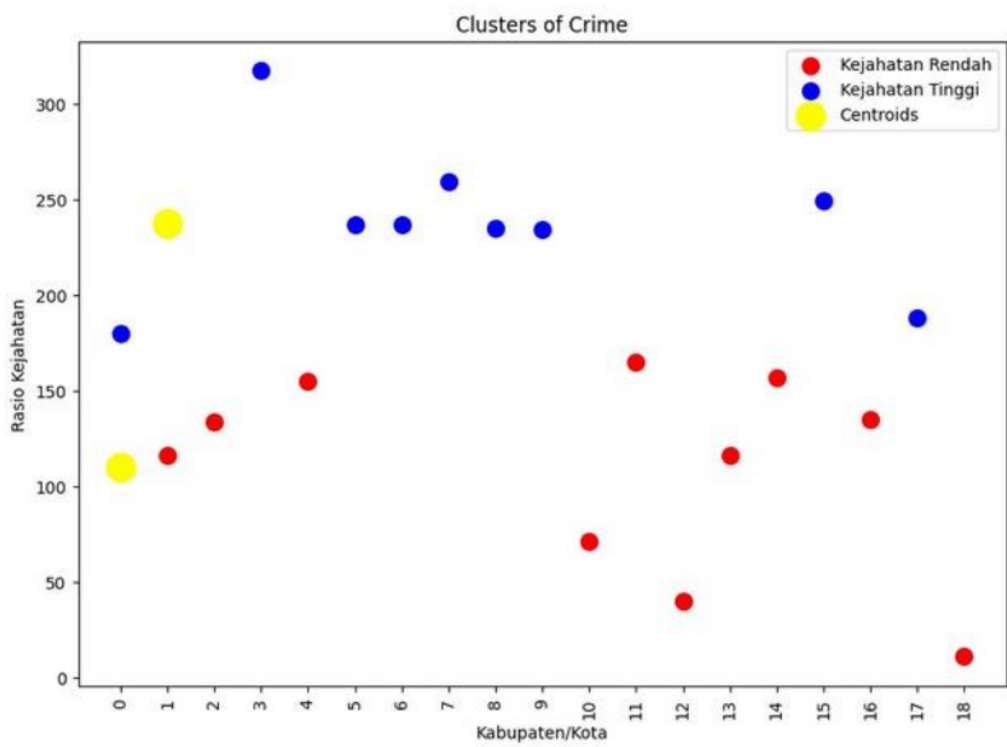
```
# Visualisasi hasil klustering
plt.figure(figsize=(10, 7))
plt.scatter(df_grouped.index[pred_y == 0], X[pred_y == 0], s=100, c='red', label = 'Kejahatan Rendah')
plt.scatter(df_grouped.index[pred_y == 1], X[pred_y == 1], s=100, c='blue', label = 'Kejahatan Tinggi')

# Plotting the centroids of the clusters
plt.scatter(range(len(kmeans.cluster_centers_)), kmeans.cluster_centers_[0], s=300, c='yellow', label = 'Centroids')

plt.title('Clusters of Crime')
plt.xlabel('Kabupaten/Kota')
plt.ylabel('Rasio Kejahatan')
plt.xticks(ticks=range(len(df_grouped.index)), labels=df_grouped.index, rotation=90) # Rotate x-axis labels for better readat
plt.legend()
plt.show()
```

Gambar 9.

Langkah-10 : Memvisualisasikan hasil clustering, dengan label Kejahatan Rendah dan Kejahatan Tinggi



IV. SIMPULAN

Algoritma K-Means yang digunakan untuk mengelompokkan data kejahatan dari 19 kabupaten/kota di Provinsi Jawa Barat. Hasilnya menunjukkan bahwa Kabupaten Bogor memiliki tingkat kejahatan tertinggi, sementara Kota Banjar memiliki tingkat kejahatan terendah. Ini menunjukkan bagaimana metode clustering dapat membantu dalam mengidentifikasi daerah-daerah dengan tingkat kejahatan yang berbeda, yang kemudian dapat digunakan oleh penegak hukum dan pembuat kebijakan untuk merancang strategi pencegahan dan penanggulangan kejahatan yang lebih efektif.

DAFTAR PUSTAKA

- [1] A. Joshi, A. S. Sabitha, and T. Choudhury, "Crime analysis using K-Means clustering," in 2017 3rd International Conference on Computational Intelligence and Networks (CINE), 2017, pp. 33–39.
- [2] R. Gustriansyah, N. Suhandi, and F. Antony, "Clustering optimization in RFM analysis based on k-means," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 18, no. 1, pp. 470–477, 2020..
- [3] N. G. P. R. Taram, I. K. G. Sukarsa, and I. G. A. M. Srinadi, "Pengelompokan tingkat kriminalitas dengan metode Agglomerative dan K-Means serta Peubah Pencirinya," *E-Jurnal Mat.*, vol. 8, no. 2, pp. 102–111, May 2019.
- [4] H. D. Tampubolon, S. Suhada, M. Safii, S. Solikhun, and D. Suhendro, "Penerapan algoritma K-Means dan K-Medoids clustering untuk mengelompokkan tindak kriminalitas berdasarkan provinsi," *J. Ilmu Komput. dan Teknol.*, vol. 2, no. 2, pp. 6–12, Nov. 2021.
- [5] S. K. Appiah, K. Wirekoh, E. N. Aidoo, S. D. Oduro, and Y. D. Arthur, "A model-based clustering of expectation-maximization and K -means algorithms in crime hotspot analysis," *Res. Math.*, vol. 9, no. 1, pp. 1–12, Dec. 2022.
- [6] E. D. Aritonang, H. S. Tambunan, J. T. Hardinata, E. Irawan, and D. Suhendro, "Penerapan data mining dalam mengelompokkan provinsi rawan kejahatan menggunakan algoritma K-Means," in KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer), 2020, pp. 35–42.
- [7] N. G. P. R. Taram, I. K. G. Sukarsa, and I. G. A. M. Srinadi, "Pengelompokan tingkat kriminalitas dengan metode Agglomerative dan K-Means serta Peubah Pencirinya," *E-Jurnal Mat.*, vol. 8, no. 2, pp. 102–111, May 2019.
- [8] A. A. Alkhaibari and Ping-Tsai Chung, "Cluster analysis for reducing city crime rates," in 2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT), 2017, pp. 1–6.
- [9] E. H. S. Atmaja, "Pengelompokan tingkat kriminalitas di kota Yogyakarta dengan menggunakan metode K-Means clustering," in Seminar Nasional Riset dan Teknologi Terapan, 2018, pp. IK7–IK15.
- [10] M. D. Simatupang and A. W. Wijayanto, "Analisis klaster berdasarkan tindakan kriminalitas di Indonesia tahun 2019," *J. Stat. Ind. dan Komputasi*, vol. 6, no. 1, pp. 10–19, 2021.
- [11] N. Azis et al., "Mapping study using the unsupervised learning clustering approach," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1088, no. 1, pp. 1–7, Feb. 2021.
- [12] B. Sivanagaleela and S. Rajesh, "Crime analysis and prediction using Fuzzy C-Means algorithm," in 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 595–599.

Optimasi Pengelompokan Tindak Kejahatan di Provinsi Jawa Barat melalui K-Means Clustering

ORIGINALITY REPORT

47%
SIMILARITY INDEX

43%
INTERNET SOURCES

27%
PUBLICATIONS

29%
STUDENT PAPERS

PRIMARY SOURCES

1 [media.neliti.com](#) 9%
Internet Source

2 Submitted to Universitas Nasional 9%
Student Paper

3 [dokterskripsi.com](#) 7%
Internet Source

4 [repository.mercubuana.ac.id](#) 2%
Internet Source

5 [ejurnal.seminar-id.com](#) 2%
Internet Source

6 Submitted to Sekolah Tinggi Sandi Negara 2%
Student Paper

7 [digilib.uinsa.ac.id](#) 1%
Internet Source

8 [journals.usm.ac.id](#) 1%
Internet Source

9 [jurnal.cic.ac.id](#) 1%
Internet Source

10

katalog.data.go.id

Internet Source

1 %

11

Adellia Meiriza, Edwar Ali, Rahmiati, Agustin.
"Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Program BPJS Ketenagakerjaan", Indonesian Journal of Computer Science, 2023

Publication

1 %

12

Rahmat Hidayat, Febri Laksana, Yulia Dewi.
"PENGELOMPOKAN JUMLAH KEKERASAN TERHADAP ANAK MENURUT KECAMATAN DI KABUPATEN BANYUMAS MENGGUNAKAN PENERAPAN ALGORITMA K-MEANS CLUSTERING", INDEXIA, 2023

Publication

1 %

13

Submitted to Universitas Maritim Raja Ali Haji

Student Paper

1 %

14

ojs.trigunadharma.ac.id

Internet Source

1 %

15

Muhammad Nur Kholis, Muhammad Nur Kholis. "PERANCANGAN APLIKASI PEMASARAN PRODUK UMKM DI DESA KENDURUAN UNTUK MENINGKATKAN JANGKAUAN PASAR DAN DAYA SAING BISNIS LOKAL", Jurnal Informatika dan Teknik Elektro Terapan, 2023

Publication

1 %

16	jurnal.peneliti.net Internet Source	1 %
17	ejournal.bsi.ac.id Internet Source	1 %
18	ejournal.itn.ac.id Internet Source	1 %
19	Alfio I. Regla, Davood Pour Yousefian Barfeh, Alexander A. Hernandez. "Clustering of Riding in Tandem Incidents using K-means: A Case Study in the Philippines", 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 2019 Publication	1 %
20	ejournal.unsri.ac.id Internet Source	1 %
21	acehraya.co.id Internet Source	1 %
22	issuu.com Internet Source	1 %
23	123dok.com Internet Source	<1 %
24	Usep Tatang Suryadi, Yana Supriatna. "SISTEM CLUSTERING TINDAK KEJAHATAN PENCURIAN DI WILAYAH JAWA BARAT MENGGUNAKAN ALGORITMA K-MEANS",	<1 %

Jurnal Teknologi dan Komunikasi STMIK Subang, 2019

Publication

Exclude quotes On

Exclude matches Off

Exclude bibliography On