

ANALISA MODEL CLUSTERING UNTUK PEMETAAN KUALITAS LULUSAN MAHASISWA BERDASARKAN DATASET TRACER STUDY

Riki Andri Yusda✉, Risnawati, Santoso, Putri Zakiyah Maharani Siregar,

Widiya Putri Nurani

Universitas Royal, Kisaran, Indonesia

Email: rikiandriyusda@gmail.com

ABSTRACT

Graduate data from the tracer study process has a very important role in assessing the quality of a university's graduates. From this data, universities can see an objective picture to measure and evaluate the curriculum, materials, and the achievement of learning competencies so far whether they are in accordance with what is expected by graduate users. This will provide input to university management in making strategies and policies to improve quality. However, the problem is that the amount of data available so far has not been maximized properly to assist management in making decisions. Data on graduates and users of existing graduates are only processed into semester and annual reports and there is no in-depth analysis. So that management does not get information that helps to improve the quality of graduates in the future. Optimization of clustering methods using elbow method with comparison of other distance formulas such as Euclidean Distance, Mahalanobis Distance and Manhattan City Distance to improve the performance of mapping results. The DBI result obtained is 1.89 for the number of 6 clusters.

Keyword: Graduate Quality, Tracer Study, Clustering, Elbow Method.

ABSTRAK

Data lulusan dari proses tracer study memiliki peranan yang sangat penting dalam menilai kualitas lulusan sebuah perguruan tinggi. Dari data ini perguruan tinggi dapat melihat gambaran dengan objektif untuk mengukur dan mengevaluasi kurikulum, materi, hingga ketercapaian kompetensi belajar selama ini apakah sesuai dengan yang diharapkan oleh pengguna lulusan. Hal ini akan memberikan masukan kepada manajemen perguruan tinggi dalam membuat strategi dan kebijakan untuk meningkatkan mutu. Akan tetapi, permasalahannya adalah banyaknya data yang ada selama ini belum dimaksimalkan dengan baik untuk membantu manajemen dalam pengambilan keputusan. Data lulusan dan pengguna lulusan yang ada hanya diolah menjadi laporan semester dan tahunan saja dan tidak ada analisis yang mendalam. Sehingga manajemen tidak mendapatkan informasi yang membantu untuk perbaikan kualitas lulusan dimasa yang akan datang. Optimasi metode clustering menggunakan elbow method dengan perbandingan formula jarak lainnya seperti Euclidean Distance, Mahalanobis Distance dan Manhattan City Distance untuk meningkatkan performance hasil pemetaan. Hasil DBI yang didapatkan senilai 1.89 untuk jumlah 6 cluster.

Kata Kunci: Kualitas Lulusan, Pelacakan Alumni, Klasterisasi, Elbow Method.

PENDAHULUAN

Kemajuan teknologi komputer telah dirasakan pada seluruh aspek kehidupan salah satunya pada dunia pendidikan. Teknologi ini banyak dimanfaatkan untuk memudahkan manusia secara individu maupun organisasi untuk memudahkan pekerjaannya. Meningkatnya pengguna teknologi komputer dengan dukungan jaringan internet dan berbagai aplikasi, menghasilkan kumpulan data yang sangat besar. Kumpulan data tersebut dapat diolah untuk menghasilkan informasi yang dapat membantu manusia dalam pengambilan keputusan. Dalam dunia pendidikan tinggi, menurut data Dirjen Pendidikan Tinggi Riset dan Teknologi pada tahun 2021

menjelaskan sebanyak 1,7 juta mahasiswa jenjang sarjana lulus setiap tahunnya. Data lulusan sangat mudah didapatkan dengan teknologi saat ini.

Data lulusan dari proses tracer study memiliki peranan yang sangat penting dalam menilai kualitas lulusan sebuah perguruan tinggi (Fenta et al., 2019)(Kalaw, 2019). Dari data ini perguruan tinggi dapat melihat gambaran dengan objektif untuk mengukur dan mengevaluasi kurikulum, materi, hingga ketercapaian kompetensi belajar selama ini apakah sesuai dengan yang diharapkan oleh pengguna lulusan. Hal ini akan memberikan masukan kepada manajemen perguruan tinggi dalam membuat strategi dan kebijakan

untuk meningkatkan mutu (Rahim & Puluhulawa, 2018)(Sagena et al., 2019).

Akan tetapi, permasalahannya adalah banyaknya data yang ada selama ini belum dimaksimalkan dengan baik untuk membantu manajemen dalam pengambilan keputusan. Misalnya data tracer study di Lembaga Pusat Layanan Karier STMIK Royal, data lulusan dan pengguna lulusan yang ada hanya diolah menjadi laporan semester dan tahunan saja dan tidak ada analisis yang mendalam. Sehingga manajemen tidak mendapatkan informasi yang membantu untuk perbaikan kualitas lulusan dimasa yang akan datang. Untuk menggali data menjadi beberapa cluster yang dapat dilakukan adalah menggunakan metode K-means clustering (Astuti et al., 2019)(Nabila et al., 2021). Metode clustering dengan algoritma K-Means dan optimasi elbow method dipilih dikarenakan metode ini memiliki kemampuan memartisi data ke dalam beberapa cluster data berdasarkan karakteristiknya sehingga dapat menghemat iterasi yang ada (Syakur et al., 2018)(Xiao et al., 2018). Melalui metode ini data lulusan akan dicluster yang menunjukkan tingkatan kualitasnya berdasarkan faktor-faktor yang mempengaruhinya. Untuk menentukan jumlah cluster optimal pada metode K-means, maka akan diterapkan elbow method. Sedangkan untuk mendapatkan hasil performance cluster yang terbaik, maka penelitian ini akan membandingkan performance beberapa jarak data yang ada seperti jarak Euclidean Distance, Mahalanobis Distance dan Manhattan City yang di terapkan pada metode k-means. Pengukuran performance cluster akan menggunakan Davies Bouildin Index (DBI).

Pendekatan pemecahan masalah dimulai dari proses persiapan data sesuai dengan terdapat tahapan metode clustering dengan pendekatan Knowledge Discovery in Database (KDD) antara lain data seleksi, data preprocessing, transformasi data (Maulida et al., 2018). Setelah data siap digunakan, maka proses selanjutnya adalah akan dilakukan analisa optimasi jumlah cluster optimal menggunakan elbow method. Performance model clustering sering kali menurun ketika memproses data outliers (pencilan) dalam kumpulan data yang mengakibatkan proses iterasi menjadi sangat panjang (Aziz, 2022)(Aprius et al., 2018). Untuk mengatasi data outliers dan iterasi yang panjang ini, peneliti menggunakan elbow method. Selanjutnya akan dilakukan perbandingan formula jarak data lainnya seperti Euclidean Distance, Mahalanobis Distance dan Manhattan City Distance untuk mendapatkan model cluster dengan performance terbaik dalam clusterisasi kualitas lulusan dari dataset tracer study. Tools yang digunakan yaitu jupyter

notebook dengan bahasa Python. Dalam pengolahan data besar, terdapat tahapan metode clustering yang dikenal dengan Knowledge Discovery in Database (KDD) (Safhi et al., 2019).

TINJAUAN PUSTAKA

K-Means Clustering

Langkah-langkah dalam metode clustering dengan algoritma k-means (Nurdiyansyah et al., 2021):

1. Menentukan jumlah cluster.
2. Menentukan nilai centroid.
3. Menghitung jarak antara titik centroid dengan titik tiap objek.
4. Pengelompokan objek dengan memperhitungkan jarak minimum objek.
5. Kembali ke tahap ke-2.

Mahalanobis Distance

Mahalanobis distance merupakan salah satu perhitungan jarak sama seperti Euclidean distance. Mahalanobis menggunakan suatu korelasi antar variabel untuk menghitung jarak. Jarak antara objek pertama dan kedua yang memiliki kemiripan yang tinggi dapat diukur menggunakan jarak ini untuk diklasifikasikan. Perbedaan antara Mahalanobis dengan Euclidean adalah pada jarak yang memperhitungkan variabel, di mana Mahalanobis memiliki bidang statistik multivariat yang luas (Handoyo et al., 2023). Perhitungan Mahalanobis diukur melalui ukuran mean dan median berdasarkan dispersi multivariat matriks kovariat. Jarak antara objek akan diukur dengan menghitung mean vector dan matriks kovarian. Penerapan jarak mahalanobis memiliki manfaat dalam klasifikasi, pengelompokan, analisis pola diskriminan, analisis komponen, pemrosesan gambar, neurocomputing, dan fisika. Adapun persamaan untuk mahalanobis distance sebagai berikut (Handoyo et al., 2023)(Berrendero et al., 2020)(Margolang et al., 2022):

$$D_M(x, \mu) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)}$$

Keterangan:

x : vektor data yang diamati

μ : vektor mean dari distribusi

S^{-1} : matriks invers dari matriks kovarian S

$(x - \mu)^T$: transpos dari vektor perbedaan antara x dan μ

Manhattan Distance

Manhattan distance merupakan metode perhitungan jarak data dengan menerapkan konsep selisih mutlak (Miftahuddin et al., 2020). Persamaan

manhattan distance dikenal juga sebagai L1 distance atau taxicab distance antara dua titik $x = (x_1, x_2, x_3, \dots, x_n)$ dan $y = (y_1, y_2, y_3, \dots, y_n)$ dalam ruang berdimensi n adalah (Miftahuddin et al., 2020)(Hartono et al., 2023)(Setiawan, 2022):

$$D_M(x, y) = \sum_{i=1}^n |x_i - y_i|$$

Keterangan:

(x,y) : jarak

x : koordinat lokasi 1

y : koordinat lokasi 2

Davies-Bouldin Index (DBI)

DBI merupakan metode yang digunakan untuk mengevaluasi hasil cluster. Validasi yang dilakukan adalah seberapa baik clustering yang sudah dilakukan dengan menghitung kuantitas dan fitur turunan dari data set. Jumlah klaster terbaik ditunjukkan dengan nilai DB yang semakin kecil (Alifah & Fauzan, 2023). Indeks validitas Davies-Bouldin Index (DBI) menghitung rata-rata nilai setiap titik pada himpunan data (Luthfi & Wijayanto, 2021). Jumlah klaster terbaik ditunjukkan dengan nilai DBI yang semakin kecil dengan rumus berikut.

$$DBI = \frac{1}{N} \sum_{i=1}^N \max_{j \neq i} \left(\frac{s_i + s_j}{d_{ij}} \right)$$

Keterangan:

N : jumlah cluster

s_i : ukuran rata-rata sebaran dalam cluster i

d_{ij} : jarak antara centroid cluster i dan j

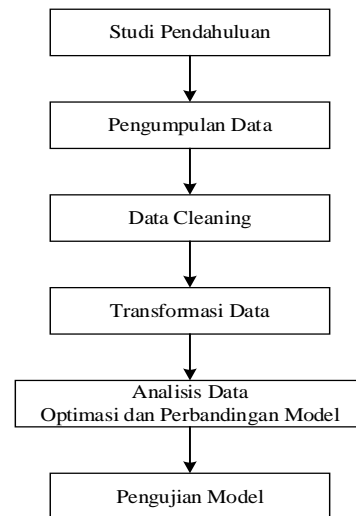
$\max_{j \neq i} \left(\frac{s_i + s_j}{d_{ij}} \right)$: rasio dispersi dan jarak antara cluster

i dan j

Semakin rendah nilai DBI, semakin baik clustering, karena ini menunjukkan cluster yang lebih compact dan lebih jauh satu sama lain.

METODE PENELITIAN

Untuk mencapai tujuan penelitian yang direncanakan, peneliti menyusun tahapan-tahapan penelitian dan pembagian tugas dari tiap tahapan sebagaimana bagan berikut.



Gambar 1. Tahapan Penelitian

1. **Studi Pendahuluan**
 Pada bagian ini akan dilakukan pengamatan lapangan, kajian pustaka, pemahaman tentang pokok persoalan.
2. **Pengumpulan Data**
 Pengumpulan data dilakukan menggunakan instrumen penelitian dan wawancara terkait lulusan dari dataset tracer study.
3. **Data Cleaning**
 Data Cleaning merupakan proses pembersihan data. Ini dilakukan untuk memperbaiki data bermasalah seperti error dan ganda.
4. **Transformasi Data**
 Sesuai dengan metode clustering, pada tahap ini akan dilakukan perubahan data dari kategorial dan verbal menjadi numerik. Ini dilakukan agar tidak terjadi missing value pada saat pemrosesan data.
5. **Analisis Data**
 Pada tahap ini akan dilakukan pengolahan data menggunakan metode clustering algoritma K-Means dengan optimasi menggunakan elbow method. Kemudian dilakukan perbandingan performance formula jarak data yaitu Euclidean, Mahalanobis Distance, dan Manhattan City untuk mendapatkan jarak yang paling optimal. Ini dilakukan dalam upaya optimasi peningkatan performance metode Clustering dalam sebuah sistem pendukung keputusan menggunakan jupyter notebook.
6. **Pengujian Model**
 Pengujian model dilakukan untuk mengetahui performance dari model yang dipilih. Pengujian menggunakan Davies-Bouldin Index (DBI) dengan menghitung rata-rata nilai setiap titik pada himpunan data. Perhitungan nilai setiap titik adalah jumlah nilai compactness yang dibagi dengan jarak

antara kedua titik pusat kluster sebagai separation. Jumlah kluster terbaik ditunjukkan dengan nilai DB yang semakin kecil. Pengujian dilakukan dengan Davies-Bouldin Index (DBI) menggunakan Jupyter Notebook dengan bahasa Python.

HASIL DAN PEMBAHASAN

Pre-processing Data

Dalam pemilihan data dari objek penelitian merupakan data kepuasan pengguna lulusan terhadap alumni dari perguruan tinggi. Dari data yang telah didapatkan, kemudian dilakukan proses normalisasi terhadap data agar dapat dilakukan pemrosesan. Berikut perbandingan data sebelum dan setelah dilakukan normalisasi data dengan menggunakan python.

| Kerjasama_Tim | Keahlian_IT | Kemampuan_Bahasa_Ingggris | Kemampuan_Berkomunikasi | Pengembangan_Diri | Kepemimpinan | Etos_Kerja |
|---------------|-------------|---------------------------|-------------------------|-------------------|--------------|-------------|
| Cukup | Baik | Cukup | Kurang | Kurang | Cukup | Cukup |
| Kurang | Cukup | Kurang | Baik | Cukup | Cukup | Sangat Baik |
| Sangat Baik | Sangat Baik | Baik | Baik | Cukup | Cukup | Sangat Baik |
| Cukup | Sangat Baik | Kurang | Kurang | Cukup | Cukup | Kurang |
| Cukup | Sangat Baik | Kurang | Sangat Baik | Kurang | Kurang | Cukup |
| Kurang | Sangat Baik | Cukup | Sangat Baik | Baik | Cukup | Baik |
| Sangat Baik | Kurang | Kurang | Baik | Kurang | Sangat Baik | Kurang |
| Sangat Baik | Sangat Baik | Cukup | Sangat Baik | Baik | Cukup | Cukup |
| Cukup | Kurang | Sangat Baik | Kurang | Kurang | Kurang | Cukup |
| Baik | Kurang | Sangat Baik | Sangat Baik | Kurang | Cukup | Baik |

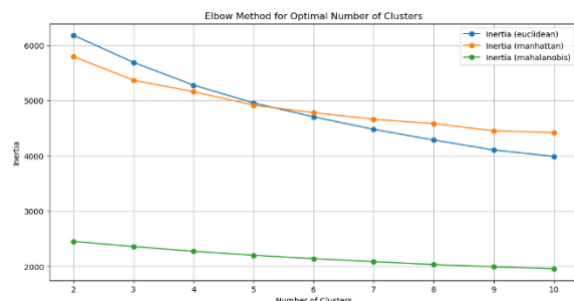
Gambar 2. Data sebelum Normalisasi Data

| Kerjasama_Tim | Keahlian_IT | Kemampuan_Bahasa_Ingggris | Kemampuan_Berkomunikasi | Pengembangan_Diri | Kepemimpinan | Etos_Kerja |
|---------------|-------------|---------------------------|-------------------------|-------------------|--------------|------------|
| 1.0 | 2.0 | 1.0 | 0.0 | 0.0 | 1.0 | 1.0 |
| 0.0 | 1.0 | 0.0 | 2.0 | 1.0 | 1.0 | 3.0 |
| 3.0 | 3.0 | 2.0 | 2.0 | 1.0 | 1.0 | 3.0 |
| 1.0 | 3.0 | 0.0 | 0.0 | 1.0 | 1.0 | 0.0 |
| 1.0 | 3.0 | 0.0 | 3.0 | 0.0 | 0.0 | 1.0 |
| 0.0 | 3.0 | 1.0 | 3.0 | 2.0 | 1.0 | 2.0 |
| 3.0 | 0.0 | 0.0 | 2.0 | 0.0 | 3.0 | 0.0 |
| 3.0 | 3.0 | 1.0 | 3.0 | 2.0 | 1.0 | 1.0 |
| 1.0 | 0.0 | 3.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| 2.0 | 0.0 | 3.0 | 3.0 | 0.0 | 1.0 | 2.0 |
| 1.0 | 3.0 | 1.0 | 2.0 | 2.0 | 3.0 | 0.0 |

Gambar 3. Hasil Normalisasi Data

Elbow Method

Setelah data ternormalisasi, langkah berikutnya adalah menentukan cluster terbaik dengan menggunakan elbow method. Ditunjukkan oleh gambar berikut.



Gambar 4. Elbow Method

Inersia adalah jumlah jarak kuadrat antara setiap titik data dan pusat kluster yang ditetapkan. Nilai inersia yang lebih rendah menunjukkan bahwa cluster lebih rapat. Jumlah Cluster (sumbu X) menunjukkan jumlah cluster (k) yang digunakan dalam algoritma K-Means. Grafik ini membandingkan nilai k yang berbeda (dari 3 hingga 10).

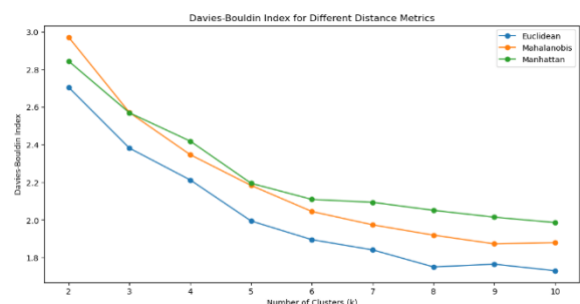
Garis yang berwarna biru adalah metrik jarak standar yang digunakan dalam K-Means. Dengan bertambahnya jumlah cluster, inersia berkurang. Namun, setelah sekitar 5 cluster, pengurangan inersia menjadi kurang jelas, yang menunjukkan bahwa peningkatan jumlah cluster lebih lanjut tidak memberikan peningkatan yang signifikan dalam kualitas pengelompokan.

Garis warna jingga menunjukkan jarak Manhattan menunjukkan pola yang mirip dengan Euclidean, dengan inersia yang menurun dengan cepat pada awalnya, kemudian melambat sekitar 5-6 cluster. Titik "siku" berada di sekitar 5-6 kluster, yang kemungkinan besar merupakan jumlah kluster yang optimal.

Garis warna hijau mewakili jarak Mahalanobis menunjukkan nilai inersia yang jauh lebih rendah dibandingkan dengan dua metrik lainnya. Hal ini dikarenakan jarak Mahalanobis memperhitungkan korelasi antar variabel, sehingga menghasilkan cluster yang lebih ringkas. Inersia relatif stabil di berbagai jumlah cluster, yang berarti bahwa menambahkan lebih banyak cluster tidak secara drastis mengurangi jumlah jarak.

Davies-Bouldin Index

Langkah selanjutnya, membandingkan setiap jarak dengan menggunakan Euclidean, Mahalanobis, dan Manhattan Distance dengan pengukuran kinerja menggunakan DBI. Sehingga di dapatkan jarak pengukuran cluster yang ditunjukkan oleh gambar berikut.



Gambar 5. Hasil Pengukuran Jarak dengan DBI

Grafik di atas mengilustrasikan Indeks Davies-Bouldin (DBI) untuk jumlah cluster (k) yang berbeda

di tiga metrik jarak: Euclidean, Mahalanobis, dan Manhattan.

Indeks Davies-Bouldin (Sumbu Y): Indeks Davies-Bouldin mengukur kualitas pengelompokan, dengan nilai yang lebih rendah mengindikasikan pemisahan dan kekompakan yang lebih baik dari cluster. Jumlah Cluster (Sumbu X): Ini menunjukkan jumlah cluster (k) yang digunakan dalam algoritma K-Means. Grafik menunjukkan nilai DBI untuk jumlah klaster mulai dari 2 hingga 10.

Pada garis Euclidean distance, Indeks Davies-Bouldin terus menurun seiring dengan bertambahnya jumlah klaster, yang mengindikasikan bahwa kualitas pengelompokan meningkat secara konsisten dengan bertambahnya jumlah klaster. DBI terendah adalah sekitar 9 cluster, menunjukkan bahwa ini mungkin merupakan jumlah cluster yang optimal dengan menggunakan metrik jarak Euclidean.

Pada garis Mahalanobis, jarak Mahalanobis menunjukkan penurunan yang lebih tajam pada DBI seiring bertambahnya jumlah cluster dari 2 ke 6, tetapi peningkatannya menjadi minimal setelah 6 cluster. DBI mencapai titik terendah setelah 6 cluster, yang mengindikasikan bahwa penambahan jumlah cluster tidak secara signifikan meningkatkan kualitas pengelompokan.

Pada garis Manhattan pada awalnya memiliki nilai DBI yang lebih tinggi dibandingkan dengan Euclidean dan Mahalanobis untuk jumlah cluster yang lebih rendah ($k = 2, 3, 4$), tetapi menjadi stabil setelah 5 cluster. Setelah 6 klaster, DBI berubah sangat sedikit, mengindikasikan bahwa meningkatkan jumlah klaster di luar titik ini tidak meningkatkan kualitas pengelompokan.

Jarak Euclidean: Metrik ini secara konsisten menunjukkan kinerja pengelompokan yang lebih baik (DBI yang lebih rendah) seiring dengan bertambahnya jumlah cluster. Jumlah cluster yang optimal mungkin sekitar 8-9, berdasarkan nilai DBI terendah.

Jarak Mahalanobis: Jarak Mahalanobis menunjukkan peningkatan yang signifikan hingga 5-6 klaster, setelah itu manfaat dari penambahan jumlah klaster akan berkurang. Jumlah cluster yang optimal untuk Mahalanobis tampaknya sekitar 6, setelah itu DBI tetap cukup konstan. Jarak Manhattan: Metrik ini meningkat dengan cepat hingga 5 klaster, tetapi mirip dengan Mahalanobis, metrik ini menunjukkan peningkatan yang minimal setelah titik ini. Jumlah cluster yang optimal untuk Manhattan mungkin juga sekitar 5-6. Hasil DBI yang dihasilkan adalah 1.89 untuk 6 cluster.

KESIMPULAN

Titik siku untuk jarak Euclidean dan Manhattan kemungkinan sekitar 5 atau 6 cluster, di mana pengurangan inersia mulai berkurang. Jarak Mahalanobis memberikan inersia yang jauh lebih rendah secara keseluruhan, yang mengindikasikan bahwa pengelompokannya lebih rapat. Namun, ini tidak menunjukkan "siku" yang jelas, yang menunjukkan bahwa jumlah cluster yang optimal lebih sulit untuk ditentukan dengan menggunakan metrik ini saja. Hal ini menunjukkan bahwa meskipun jarak Euclidean dan Manhattan dapat bekerja dengan cara yang sama, Mahalanobis dapat memberikan pendekatan pengelompokan yang lebih kuat, tergantung pada struktur dan korelasi dalam data Anda.

Jarak Euclidean menunjukkan kinerja pengelompokan yang lebih baik secara keseluruhan untuk jumlah cluster yang lebih besar, dan 8-9 cluster mungkin optimal. Jarak Mahalanobis dan Manhattan menunjukkan tren yang sama, dengan sekitar 5-6 klaster yang optimal sebelum nilai DBI stabil, yang menunjukkan berkurangnya hasil dari penambahan lebih banyak klaster.

Dari hasil pengukuran dengan Davies-Bouldin Indeks dihasilkan nilai sebesar 1.89 untuk 6 cluster.

DISEMINASI

Artikel ini telah diseminasikan pada Seminar Nasional Teknologi Informasi dan Komunikasi (SEMNASTIK) APTIKOM Tahun 2024 yang diselenggarakan oleh Universitas Methodist Indonesia pada tanggal 24-26 Oktober 2024.

DAFTAR PUSTAKA

- Alifah, R. F. N., & Fauzan, Abd. C. (2023). Implementasi Algoritma K-Means Clustering Berbasis Jarak Manhattan untuk Klasterisasi Konsentrasi Bidang Mahasiswa. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 5(1), 31–41. <https://doi.org/10.28926/ilkomnika.v5i1.542>
- Aprius, S., #1, S., #2, A. I., & Sedyono, E. (2018). *Metode K-Means Clustering dengan Atribut RFM untuk Mempertahankan Pelanggan*. 4, 433. <https://doi.org/10.28932/jutisi.v4i3.878>
- Astuti, D., Rahmat Iskandar, A., & Febrianti, A. (2019). *Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering*. 1(2), 60–072. <https://doi.org/10.20895/INISTA.V1I2>
- Aziz, S. (2022). *Implementation of Self Organizing Map Algorithm to Identify the Grouping Patterns of Family Welfare Level in Siak Regency*. 2(2).
- Berrendero, J. R., Bueno-Larraz, B., & Cuevas, A. (2020). On Mahalanobis Distance in Functional

- Settings. In *Journal of Machine Learning Research* (Vol. 21).
<http://jmlr.org/papers/v21/18-156.html>.
- Fenta, H. M., Asnakew, Z. S., Debele, P. K., Nigatu, S. T., & Muhaba, A. M. (2019). Analysis of supply side factors influencing employability of new graduates: A tracer study of Bahir Dar University graduates. *Journal of Teaching and Learning for Graduate Employability*, 10(2), 67–85. <https://ojs.deakin.edu.au/index.php/jtlge/>
- Handoyo, R., Erny Herwindiati, D., Sutrisno, T., Tarumanagara, U., & Kunci, K. (2023). Klasifikasi Buah Busuk Dan Buah Matang Berdasarkan Data Image Menggunakan Mahalanobis Distance. *Humantech: Jurnal Ilmiah Multi Disiplin Indonesia*, 2(4), 640-654.
- Hartono, B., Eniyati, S., & Hadiono, K. (2023). Perbandingan Metode Perhitungan Jarak pada Nilai Centroid dan Pengelompokan Data Menggunakan K-Means Clustering. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(3), 503. <https://doi.org/10.30865/json.v4i3.6021>
- Kalaw, M. T. B. (2019). Tracer study of bachelor of science in mathematics. *International Journal of Evaluation and Research in Education*, 8(3), 537–548.
<https://doi.org/10.11591/ijere.v8i3.17343>
- Luthfi, E., & Wijayanto, A. W. (2021). Analisis perbandingan metode hirarchical, k-means, dan k-medoids clustering dalam pengelompokan indeks pembangunan manusia Indonesia. *INOVASI*, 17(4), 761–773.
<https://doi.org/10.30872/jinv.v17i4.10106>
- Margolang, K. F., Siregar, M. M., Riyadi, S., & Situmorang, Z. (2022). Analisa Distance Metric Algoritma K-Nearest Neighbor Pada Klasifikasi Kredit Macet. *Journal of Information System Research (JOSH)*, 3(2), 118–124.
<https://doi.org/10.47065/josh.v3i2.1262>
- Maulida, L., Studi, P., Informatika, M., Bsi, A., Komplek, T., Xiv-C11, S., Letnan, J., & Banten, S. (2018). Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means. *Jurnal Informatika Sunan Kalijaga*, 2(3), 167–174.
- Miftahuddin, Y., Umaroh, S., & Karim, F. R. (2020). Perbandingan Metode Perhitungan Jarak Euclidean, Haversine, Dan Manhattan Dalam Penentuan Posisi Karyawan. *Jurnal Tekno Insentif*, 14(2), 69–77.
<https://doi.org/10.36787/jti.v14i2.270>
- Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(2), 100.
<http://jim.teknokrat.ac.id/index.php/JTSI>
- Nurdiyansyah, F., Akbar, I., History, A., & Author, C. (2021). *Jurnal Teknologi dan Manajemen Informatika Implementasi Algoritma K-Means untuk Menentukan Persediaan Barang pada Poultry Shop Article Info ABSTRACT*. 7(2), 86–94.
<http://http://jurnal.unmer.ac.id/index.php/jtmi>
- Rahim, M., & Puluhalawa, M. (2018). Tracer Study Lulusan Program Studi Bimbingan dan Konseling Universitas Negeri Gorontalo. *Jurnal Pendidikan (Teori Dan Praktik)*, 2(2), 121.
<https://doi.org/10.26740/jp.v2n2.p121-127>
- Safhi, H. M., Frikh, B., & Ouhbi, B. (2019). Assessing reliability of Big Data Knowledge Discovery process. *Procedia Computer Science*, 148, 30–36. <https://doi.org/10.1016/j.procs.2019.01.005>
- Sagena, U. W., Pamilih Widagdo, P., Pranoto, H., & Zahrotun Kamila, V. (2019). *Proceeding Indonesia Career Center Network Summit IV e*. 17–18. https://perkasa.unmul.ac.id/tracer_study/.
- Setiawan, A. (2022). Perbandingan Penggunaan Jarak Manhattan, Jarak Euclid, dan Jarak Minkowski dalam Klasifikasi Menggunakan Metode KNN pada Data Iris. *Jurnal Sains Dan Edukasi Sains*, 5(1), 28–37.
<https://doi.org/10.24246/juses.v5i1p28-37>
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster. *IOP Conference Series: Materials Science and Engineering*, 336(1).
<https://doi.org/10.1088/1757-899X/336/1/012017>
- Xiao, B., Wang, Z., Liu, Q., & Liu, X. (2018). SMK-means: An Improved Mini Batch K-means Algorithm Based on Mapreduce with Big Data. In *Journal name* (Vol. 1, Issue 1).
www.techscience.com/xxx