

# Data Mining Manhattan Distance dan Euclidean Distance pada Algoritma X-Means dalam Klasifikasi Minat dan Bakat Siswa

<sup>1</sup>\*Mayang Mughnyanti, <sup>2</sup>Subhan Hafiz Nanda Ginting,  
<sup>1,2</sup>Universitas Battuta  
Medan, Indonesia

<sup>1</sup>mayangmughnyanti6614@gmail.com, <sup>2</sup>subhanhafiz16@gmail.com

\*Penulis Korespondensi

Diajukan : 25/01/2023

Diterima : 31/01/2023

Dipublikasi : 31/01/2023

## ABSTRAK

Algoritma X-Means merupakan algoritma yang digunakan untuk pengelompokan data. Algoritma x means merupakan pengembangan dari k-means. X-means clustering digunakan untuk menyelesaikan salah satunya kelemahan utama dari K-means clustering, yaitu perlunya pengetahuan sebelumnya tentang jumlah cluster (K). Dalam penelitian ini Data mining yang dibangun dapat membantu dan memberikan kemudahan dalam proses klasifikasi minat dan bakat siswa. *Setting* variabel penilaian dilakukan melalui sistem. Jadi, jika ada perubahan nilai maksimal dan minimal variabel, admin tinggal mengupdate melalui sistem tanpa harus merubah programnya. Berdasarkan penelitian yang dilakukan, dapat di ketahui bahwa dengan kumpulan nilai ujian yang di lakukan siswa, dapat di ketahui seberapa minat siswa berdasarkan hasil cluster 1 dengan bidang ilmu Bidang olahraga Bidang kedokteran Bidang teknik. Sedangkan pada cluster 2 Bidang hewani, Bidang seni, desain dan media, Bidang Pendidikan, Bidang olahraga, Bidang agama dan filsafat, Bidang bahasa.

**Kata Kunci:** Manhattan Distance, Euclidean Distance, X-Means, Minat dan Bakat Siswa

## I. PENDAHULUAN

Pendidikan memiliki kedudukan penting dalam meningkatkan kualitas sumber daya manusia guna membangun suatu kemajuan bangsa, karena pada hakekatnya pendidikan adalah sebuah proses untuk mengembangkan kepribadian diri menjadi pribadi yang berilmu dan berkarakter. Dalam upaya mencerdaskan suatu bangsa, pemerintah selalu berusaha dengan optimal dalam memberikan pengarahan dan penyuluhan kepada para guru sejak dari jenjang pendidikan keluarga, sekolah, serta lingkungan masyarakat agar selalu memperhatikan bakat anak didik untuk dapat mencapai tujuan utama dari sebuah pendidikan (Saputri & Nurrus, 2021).

Perkembangan minat peserta didik pada proses belajar itu sangat penting, jika peserta didik tidak memiliki minat terhadap hal yang ada di hadapannya maka peserta didik tidak bisa menguasai hal tersebut (Anggraini et al., n.d.).

Menurut Shaleh Abdul Rahman dalam bukunya psikologi suatu pengantar dalam perspektif islam, menjelaskan bahwa minat adalah suatu kecenderungan untuk memberikan perhatian dan bertindak terhadap orang, aktivitas atau situasi yang menjadi objek dari minat tersebut dengan disertai perasaan atau gembira (Anggraini et al., 2020).

Bakat Menurut William B. Michael (suryabrata 1995) Bakat merupakan kapasitas pada diri seseorang dalam melakukan tugasnya dan melakukan dengan pengaruh dan latihan yang dijalaninya.

Menurut Bigham (1968) Bakat sebagai kondisi atau kemampuan yang dimiliki seseorang yang memungkinkan dengan suatu latihan khusus dapat memperoleh suatu kecakapan, pengetahuan dan keterampilan khusus. Untuk klasifikasi minat dan bakat anak maka diperlukan pengelompokan menggunakan metode X-means.

Adapun tujuan dari clustering adalah agar objek-objek (data) dalam suatu kelompok yang sama (terkait) satu sama lain dan berbeda (tidak terkait) objek-objek dalam kelompok lain. Semakin besar kesamaan (homogenitas) dalam suatu kelompok dan semakin besar perbedaan antar kelompok, semakin baik atau lebih jelas pengelompokan. Salah satu algoritma yang dapat digunakan dalam pengelompokan adalah X-Means (P H Putra et al., 2022).

Algoritma X-Means merupakan pengembangan dari K-Means. Kelemahan X-Means adalah pada penentuan matriks jarak, matriks jarak merupakan faktor penting yang bergantung pada X-Means kumpulan data algoritma. Nilai matriks jarak yang dihasilkan akan mempengaruhi kinerja algoritma. Itu jarak antara dua titik data ditentukan dengan perhitungan matriks jarak dimana Euclidean Distance adalah fungsi matriks jarak yang paling banyak digunakan. (Infokum, 2021). Ada beberapa jenis jarak fungsi matriks selain Euclidean Distance yaitu Manhattan Distance, Miskowski Distance, Jarak Canberra, Jarak Braycurtis, Chi-Square dan lainnya (Infokum, 2021).

## II. STUDI LITERATUR

### 2.1. Penelitian Terdahulu

Penelitian (Saputri & Nurrus, 2021) Bakat adalah potensi yang dimiliki oleh peserta didik sejak lahir, dan kemampuan dasar yang dimiliki untuk belajar dengan tempo yang relatif cepat namun hasilnya justru lebih baik dibandingkan dengan orang lain. Minat adalah ketertarikan seseorang kepada suatu kegiatan yang ia senangi atau dorongan yang kuat dalam melakukan segala sesuatu yang menjadi keinginannya.

Penelitian (Bogucharskiy & Mashtalir, 2015) Salah satu cara untuk memecahkan sebagian kesenjangan semantik antara fitur visual tingkat rendah dan konsep manusia tingkat tinggi adalah dengan menganalisis sifat spasial bagian gambar yang diinduksi oleh segmentasi. Fuzzy modifikasi dari X-means clustering telah diusulkan dalam aplikasi untuk masalah gambar segmentasi. Kekhususan modifikasi yang diusulkan terdiri dari kemampuan untuk memproses sinyal matriks di bawah tumpang tindih kelas dan tidak adanya informasi apriori tentang jumlah cluster yang secara otomatis ditentukan selama informasi pengolahan. Implementasi algoritma dari metode ini adalah ditandai dengan kesederhanaan komputasi dan tinggi pertunjukan.

Penelitian (Kim et al., 2017) Makalah ini memperkenalkan pengelompokan yang dimodifikasi metode yang diturunkan dari kombinasi mean-shift dan x-mean konsep pengelompokan untuk mengelompokkan video dalam unit dasar adegan. Metode pengelompokan yang diusulkan memperhitungkan perbedaan waktu gambar untuk mengelompokkannya ke dalam sejumlah grup yang tidak disetel. Itu kinerja metode yang diusulkan dibandingkan dengan metode pengelompokan meanshift dan x-means. Metode yang diperkenalkan adalah efisien dalam waktu dan perhitungan serta dapat mengelompok secara terpisah gambar dengan fitur yang sama atau mirip tetapi dengan perbedaan waktu.

Dari hasil eksperimen yang telah dilakukan dapat diketahui bahwa dengan menggunakan algoritme x-means dapat mengetahui kelompok data terhadap tingkat pemahaman kurikulum KKNi menggunakan tools rapidminer. Hasil analisis menunjukkan kecenderungan yang berbeda terhadap pemahaman kurikulum KKNi di STMIK IKMI Cirebon, khususnya terhadap parameter know-how dan affection. Perlu adanya sosialisasi kembali dan peningkatan cara penyampaian kurikulum KKNi agar tingkat pemahaman dilingkungan mahasiswa dapat ditingkatkan (Anwar et al., 2019).

### 2.2. Data Mining

Secara sederhana dapat dipahami bahwa data mining atau dikenal juga dengan istilah

knowledge discovery in database (KDD) adalah serangkaian proses yang bertujuan untuk mengekstraksi pola-pola penting atau menarik dari sejumlah data berukuran sangat besar yang tidak dapat dikenali secara manual. Data mining merupakan bagian yang terintegrasi dari penemuan pengetahuan dalam database yang merupakan proses dengan urutan sebagai berikut (Supoyo, 2022).

Data Mining tidak akan lengkap tanpa mengacu pada penggunaan historis istilahistilah seperti "data mining", Pada tahun 1960-an, ketika komputer semakin banyak diterapkan pada masalah analisis data, tercatat bahwa jika Anda mencari cukup lama, Anda selalu dapat menemukan beberapa model yang sesuai dengan kumpulan data dengan baik (Mamok Andri Senubekti) (Klasifikasi et al., 2022).

Data mining disebut juga dengan knowledge-discovery in database (KDD) ataupun pattern recognition. Istilah KDD atau disebut penemuan pengetahuan data karena tujuan utama data mining adalah untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna. Sedangkan istilah pattern recognition atau disebut pengenalan pola mempunyai tujuan pengetahuan yang akan digali dari dalam bongkahan data yang sedang dihadapi (Purwa Hasan Putra et al., 2022).

### 2.3. X-Means

X-means clustering digunakan untuk mengatasi salah satu kelemahan utama dari K-means clustering perlu pengetahuan sebelumnya tentang jumlah cluster ( $K$ ). Dalam metode ini, nilai sebenarnya dari  $K$  diperkirakan dengan cara yang tidak dipantau dan hanya berdasarkan kumpulan data.  $K_{max}$  dan  $K_{min}$  sebagai batas atas dan bawah untuk kemungkinan nilai  $X$ . Pada langkah pertama X-means pengelompokan, ketahuilah bahwa pada saat ini  $X = X_{min}$ , X-berarti menemukan struktur awal dan pusat massa. Pada langkah selanjutnya, setiap cluster dalam struktur estimasi diperlakukan sebagai cluster induk, yang mana dapat dibagi menjadi dua kelompok (Mughnyanti et al., 2020).

Algoritma X-Means merupakan pengembangan dari algoritme K-Means dimana dalam XMeans terdiri dari beberapa operasi berulang hingga eksekusi berakhir serta pada algoritma ini mengoptimalkan nilai Bayesian Information Criterion (BIC). Pelleg & Moore menguraikan algoritma X-Means yang ditulis sebagai berikut (Riza et al., 2022).

1. Menentukan range cluster  $K(K_{min}, K_{max})$ .
2. Menginisiasi nilai  $K = K_{min}$ .
3. Menjalankan K-Means hingga konvergen (anggota cluster tidak mengalami perubahan letak cluster).
4. Memperbaiki struktur, langkah ini dimulai dengan memecah setiap centroid hasil langkah 3 menjadi dua children dalam arah yang berlawanan di sepanjang vektor yang dipilih secara acak. Setelah itu menjalankan K-Means secara lokal di dalam setiap cluster untuk dua cluster. Keputusan masing-masing pusat cluster sendiri dengan membandingkan nilai-nilai BIC.
5. Memperbaharui nilai  $K$ , jika  $K > K_{max}$  maka proses berhenti dan melaporkan struktur terbaik yang ditemukan selama pencarian, jika tidak kembali ke langkah 3.

### 2.4. Klasifikasi

Menurut KBBI, klasifikasi adalah penyusunan bersistem dalam kelompok atau golongan menurut kaidah atau standar yang ditetapkan. Sedangkan pengertian secara umum klasifikasi adalah suatu kegiatan dengan mengelompokkan benda yang memiliki beberapa ciri yang sama dan memisahkan benda yang tidak sama (Azis et al., 2021).

### 2.5. Euclidean Distance

Euclidean distance adalah perhitungan untuk mengukur jarak dua titik yang mempelajari hubungan antara sudut dan jarak. Dalam bahasa matematika euclidean distance seperti perhitungan Pythagoras yang digunakan untuk mengukur dua titik dalam satu dimensi. Untuk menghitung euclidean distance menggunakan persamaan 6 (Pribadi et al., 2022).

$$d = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \dots\dots\dots(6)$$

Keterangan :

- d = Jarak
- x1 = Koordinat latitude 1
- x2 = Koordinat latitude 2
- y1 = Koordinat longitude 1
- y2 = Koordinat longitude 2

**2.6. Manhattan Distance**

Manhattan distance adalah metode perhitungan pada ruang jarak dengan menerapkan konsep selisih mutlak. Untuk menghitung manhattan distance menggunakan persamaan 7 (Pangestu & Fitriani, 2022).

$$d(x, y) = \sum_{i=1}^n |xi - yi| \dots\dots\dots (7)$$

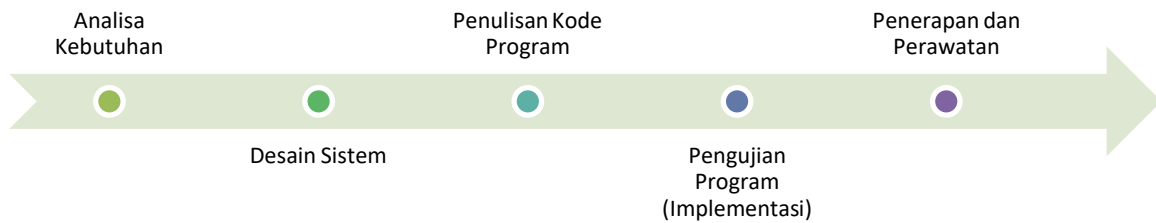
Keterangan :

- d(x, y) = Jarak
- x = Koordinat lokasi 1
- y = Koordinat lokasi 2

**III. METODE**

Metodologi dan kerangka kerja penelitian digunakan untuk menggambarkan langkah-langkah yang akan diterapkan dalam melakukan penelitian.

**Kerangka Kerja Penelitian**



**Gambar 1. Kerangka Kerja Penelitian**

Berdasarkan gambar 1 penulis dapat menjelaskan beberapa kerangka kerja yang akan dilakukan dalam penelitian ini, yaitu sebagai berikut :

1. Mempelajari Literatur  
 Pada penelitian ini dipelajari literatur yang berhubungan dengan permasalahan. Kemudian literatur yang dipelajari diseleksi untuk dapat ditentukan literatur mana yang akan digunakan dalam penelitian. Sumber literatur didapatkan dari perpustakaan, jurnal, artikel dan konsep-konsep lain yang mendukung dalam menyelesaikan sistem yang akan dibangun termasuk referensi.
2. Pengumpulan Data  
 Dalam melakukan penelitian ini, pengumpulan data dan informasi pada tahap ini dilakukan untuk mengetahui mengenai sistem yang diteliti. Dari data dan informasi yang dikumpulkan akan didapat data untuk pendukung penelitian serta pengumpulan data dilakukan untuk mengetahui kebutuhan dari pengguna. Metode yang digunakan penulis untuk pengumpulan data adalah sebagai berikut :
  - a. Observasi  
 Observasi berguna untuk melakukan pengumpulan data dan observasi dengan langsung terjun kelapangan pada pihak-pihak yang terkait dalam menyelesaikan penelitian ini dimana informasi dan materi akan diperoleh sebagai bahan dari rancang bangun sistem.
  - b. Wawancara

Melakukan wawancara pada pihak yang berkaitan dengan alur permasalahan. Wawancara ini dilakukan untuk mendapatkan bahan penulisan dan penjelasan pengamatan yang dilakukan.

3. **Analisa Kebutuhan**  
Analisis dapat didefinisikan sebagai penguraian dari suatu sistem informasi yang utuh kedalam bagian-bagian komponennya. Analisa kebutuhan ini bertujuan untuk mengetahui apa saja yang dibutuhkan dalam perancangan sistem yang di bangun dan mengetahui kebutuhan-kebutuhan pendukung dari perancangan sistem.
4. **Desain dan Perancangan Sistem**  
Kegiatan desain sistem dilakukan untuk sebagai awal dari perancangan sistem yang akan dibangun sesuai kebutuhan. Dan pada tahap ini dilakukan perancangan antar muka terhadap sistem yang akan dibuat.
5. **Implementasi Sistem**  
Implementasi sistem dilakukan sesuai desain dan rancangan antar muka aplikasi yang akan dibangun. Pada tahap ini melakukan pengkodean atau pembuatan program sehingga sistem yang dirancang dapat digunakan oleh pengguna.
6. **Pengujian Sistem**  
Pengujian sistem dilakukan untuk mengetahui uji kelayakan sistem yang telah dibangun sesuai yang diharapkan dan dengan dilakukannya pengujian dapat mengetahui kelemahan serta kelebihan dari sistem yang dirancang sehingga dapat dilakukan perbaikan pada tahap selanjutnya.
7. **Maintenance**  
Pemeliharaan Sistem.

#### IV. HASIL DAN PEMBAHASAN

Adapun penelitian ini, menggunakan algoritma X-Means sebagai metode klasifikasi data set iris dan menggunakan algoritma euclidean distance dan manhattan distance sebagai metode untuk menghitung jarak antara data tes dan data uji pada algoritma X-Means. Dan akan dianalisis dengan algoritma X-Means menggunakan metode jarak Euclidean Distance dengan perbandingan metode jarak Manhattan Distance.

- a. Tentukan kelompokkan data yang dekat dengan centroid. Adapun ditunjukkan sebagai berikut.

| Cluster 1         |      |
|-------------------|------|
| Bidang olahraga   | No.3 |
| Bidang kedokteran | No.2 |
| Bidang teknik     | No.1 |

| Cluster 2                     |              |
|-------------------------------|--------------|
| Bidang hewani                 | No.10        |
| Bidang seni, desain dan media | No.8<br>No.7 |
| Bidang Pendidikan             | No.9         |
| Bidang olahraga               | No.6         |
| Bidang agama dan filsafat     | No.4<br>No.5 |
| Bidang bahasa                 |              |
| Bidang ekonomi                |              |

- b. Hitunglah mean x dan mean y dari cluster yang terbentuk. Adapun ditunjukkan sebagai berikut.

$$\begin{aligned}
 &= \frac{x_3 + x_2 + x_1}{\text{jumlah } x} \\
 &= \frac{4.7 + 4.9 + 5.1}{3} = 4.9 \\
 &= \frac{3.2 + 3 + 3.5}{3} = 3.233 \\
 &= \frac{1.3 + 1.4 + 1.4}{3} = 1.366 \\
 &= \frac{0.2 + 0.2 + 0.2}{3} = 0.2
 \end{aligned}$$

Berikut merupakan hasil dari perhitungan untuk mencari nilai mean x dan mean y berdasarkan klasifikasi data titik centroid cluster 1. Ditunjukkan pada tabel 3.

| Cluster 1         |      | Mean X1 | Mean X2     | Mean Y1     | Mean Y2 |
|-------------------|------|---------|-------------|-------------|---------|
| Bidang olahraga   | No.3 | 4.9     | 3.233333333 | 1.366666667 | 0.2     |
| Bidang kedokteran | No.2 |         |             |             |         |
| Bidang teknik     | No.1 |         |             |             |         |

$$\begin{aligned}
 &= \frac{y_{10} + y_8 + y_7 + y_9 + y_6 + y_4 + y_5}{\text{jumlah } y} \\
 &= \frac{6.3 + 5.8 + 6.3 + 7.1 + 6.9 + 7 + 6.4}{7} = 6.542 \\
 &= \frac{2.9 + 2.7 + 3.3 + 3 + 3.1 + 3.2 + 3.2}{7} = 3.057 \\
 &= \frac{5.6 + 5.1 + 6 + 5.9 + 4.9 + 4.7 + 4.5}{7} = 5.242 \\
 &= \frac{1.8 + 1.9 + 2.5 + 2.1 + 1.5 + 1.4 + 1.5}{7} = 4.708
 \end{aligned}$$

Berikut merupakan hasil dari perhitungan untuk mencari nilai mean x dan mean y berdasarkan klasifikasi data titik centroid cluster 2. Ditunjukkan pada tabel 3.

| Cluster 2                     |       | Mean X1 | Mean X2 | Mean Y1 | Mean Y2 |
|-------------------------------|-------|---------|---------|---------|---------|
| Bidang hewani                 | No.10 | 6.542   | 3.057   | 5.242   | 1.8142  |
| Bidang seni, desain dan media | No.8  |         |         |         |         |
| Bidang Pendidikan             | No.7  |         |         |         |         |
| Bidang olahraga               | No.9  |         |         |         |         |
| Bidang agama dan filsafat     | No.6  |         |         |         |         |
| Bidang bahasa                 | No.4  |         |         |         |         |
| Bidang ekonomi                | No.5  |         |         |         |         |

**Tabel 3.5** Kelompok Cluster Setiap Data

| No | Nama Item        | Kelompok Cluster |
|----|------------------|------------------|
| 1  | Prayogi Wiranata | Cluster 1        |
| 2  | Siti Mardiyah    | Cluster 1        |
| 3  | Salsabila        | Cluster 1        |
| 4  | Novi Raphita     | Cluster 2        |
| 5  | Dita Amelia      | Cluster 2        |
| 6  | Hifzan Rahim     | Cluster 2        |
| 7  | Imam Ahmad       | Cluster 2        |

|    |              |           |
|----|--------------|-----------|
|    | Farhan       |           |
| 8  | Novi Raphita | Cluster 2 |
| 9  | Ayu Andari   | Cluster 2 |
| 10 | Ali Subur    | Cluster 2 |

Setelah proses cluster, dari tabel 3 dapat dilihat pengelompokkan cluster setiap data berdasarkan titik pusat cluster. Adapun berikut merupakan hasil dari perhitungan iterasi ke-1, dimana Data ke 2 dan 8 menjadi titik pusat cluster. Untuk iterasi selanjutnya dihitung dengan tahap yang sama sampai ditemukannya konvergen setiap data yang di klasifikasi.

## V. KESIMPULAN

Adapun kesimpulan dari penerapan metode x-means dalam klasifikasi minat dan bakat siswa yang penulis bangun sebagai berikut :

Data mining yang dibangun dapat membantu dan memberikan kemudahan dalam proses klasifikasi minat dan bakat siswa. *Setting* variabel penilaian dilakukan melalui sistem. Jadi, jika ada perubahan nilai maksimal dan minimal variabel, admin tinggal mengupdate melalui sistem tanpa harus merubah programnya. Berdasarkan penelitian yang dilakukan, dapat di ketahui bahwa dengan kumpulan nilai ujian yang di lakukan siswa, dapat di ketahui seberapa minat siswa berdasarkan hasil cluster 1 dengan bidang ilmu Bidang olahraga Bidang kedokteran Bidang teknik. Sedangkan pada cluster 2 Bidang hewani, Bidang seni, desain dan media, Bidang Pendidikan, Bidang olahraga, Bidang agama dan filsafat, Bidang bahasa.

## VI. UCAPAN TERIMA KASIH

Terima kasih kepada tempat mengabdikan kami di Universitas Battuta yang sudah memberikan motivasi terhadap kami dan terima kasih kepada keluarga kami yang paling kami sayangi.

## VII. REFERENSI

- Angraini, I. A., Utami, W. D., & Rahma, S. B. (2020). Analisis Minat dan Bakat Peserta didik terhadap Pembelajaran. *Jurnal Pendidikan Dan Pembelajaran Dasar*, 7(1), 23–28. <http://ejournal.radenintan.ac.id/index.php/terampil/index%0ATerampil>:
- Angraini, I. A., Utami, W. D., Rahma, S. B., & Tangerang, U. M. (n.d.). *Mengidentifikasi Minat Bakat Siswa Sejak 2*, 161–169.
- Anwar, S., Nuris, N. D., & Wijaya, Y. A. (2019). Pengelompokkan Tingkat Pemahaman Kurikulum Berbasis KKNI Menggunakan Metode X-Means Clustering. *Jurnal Informatika: Jurnal Pengembangan IT Poltek Tegal*, 04(2), 187–190. <https://doi.org/10.30591/jpit.v4i2-2.1869>
- Azis, A., Putra Pamungkas, D., & Setiawan, A. B. (2021). Analisa Perbandingan Algoritma Euclidean Dan Manhattan Distance Dalam Identifikasi Wajah. *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 5(1), 219–224. <https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/951>
- Bogucharskiy, S., & Mashtalir, V. (2015). Image Segmentation via X-Means under Overlapping Classes. *International Scientific and Technical Conference on Computer Sciences and Information Technologies, September*, 45–47. <https://doi.org/10.1109/STC-CSIT.2015.7325427>
- Infokum, J. (2021). *APPLICATION OF X-MEANS METHOD FOR GROUPING EARLY*. 10(1), 50–55.
- Kim, N., Park, H., Choi, J. K., & Yang, J. (2017). Time gap accounted video scene segmentation with modified mean-shift X-means clustering. *2017 IEEE 6th Global Conference on Consumer Electronics, GCCE 2017, 2017-January (Gcce)*, 1–2. <https://doi.org/10.1109/GCCE.2017.8229332>

- 
- Klasifikasi, P., Data, D. A. N., Dengan, M., & Algoritma, C. (2022). *Volume 16 Nomor 2 , Juli 2022 JURNAL NUANSA INFORMATIKA Volume 16 Nomor 2 , Juli 2022*. 16, 87–93.
- Mughnyanti, M., Efendi, S., & Zarlis, M. (2020). Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation. *IOP Conference Series: Materials Science and Engineering*, 725(1), 0–6. <https://doi.org/10.1088/1757-899X/725/1/012128>
- Pangestu, M. S., & Fitriani, M. A. (2022). Perbandingan Perhitungan Jarak Euclidean Distance, Manhattan Distance, dan Cosine Similarity dalam Pengelompokan Data Bibit Padi Menggunakan Algoritma K-Means. *Sainteks*, 19(2), 141. <https://doi.org/10.30595/sainteks.v19i2.14495>
- Pribadi, W. W., Yunus, A., & Sartika Wiguna, A. (2022). Perbandingan Metode K-Means Euclidean Distance Dan Manhattan Distance Pada Penentuan Zonasi Covid-19 Di Kabupaten Malang. *Jurnal Mahasiswa Teknik Informatika*, 6(2), 493–500.
- Putra, P H, Hasibuan, A., & Marpaung, E. A. (2022). Analisis Klasifikasi Metode X-Means Pada Minat dan Bakat Anak Dimasa Pandemi. *SITEKIN: Jurnal Sains ...*, 19(2), 424–429. <http://ejournal.uin-suska.ac.id/index.php/sitekin/article/view/17889>
- Putra, Purwa Hasan, Novelan, M. S., & Rizki, M. (2022). *ANALYSIS K-NEAREST NEIGHBOR METHOD IN CLASSIFICATION OF VEGETABLE QUALITY BASED ON COLOR*. 3(2), 126–132.
- Riza, A. A., Retno, D., Saputro, S., & Maret, S. (2022). Clustering Data Numerik Menggunakan Algoritme X-Means. *Seminar Nasional Matematika, Geometri, Statistika, Dan Komputasi*.
- Saputri, N., & Nurris, S. (2021). Pengembangan Minat dan Bakat Peserta Didik Melalui Kegiatan Ekstrakurikuler. *Jurnal Bimbingan Konseling Islam*, 2(2), 172–187.
- Supoyo, A. (2022). *Bianglala Informatika Analisis Data Mining Untuk Memprediksi Lama Perawatan Pasien Covid-19 Di DIY*. 10(1), 2022.