

Perbandingan Metode Euclidean Distance dan Haversine Distance pada Aplikasi Sistem PPDB dan algoritma K-Means Untuk Menentukan Kebijakan Peraturan Zonasi

Mustofa Kamal Syarifudin¹, Ratih Titi Komala Sari²

^{1,2} Informatika, Universitas Nasional, Indonesia

Article Info

Article history:

Received Oct 15, 2022

Revised Dec 03, 2022

Accepted Dec 09, 2022

Keywords:

euclidean
haversine
K-Means
PPDB
Jalur Zonasi

ABSTRACT

At this time, registration for public schools is straightforward to do online with a device via a web browser without the need to install an application that can ease the device's performance. Still, the government regulates it through a zoning policy that makes students register for schools close to their homes. This study examines and compares which Euclidean and Haversine algorithms are more accurate to implement in making an application that determines the distance between the school and the student's house. Then the school will decide which students can be accepted using the K-Means algorithm, as has been done by SMPN 1 Tigaraksa, which results that the haversine algorithm has an average accuracy rate of 99.71%, an average error of 0.29% with an average distance difference of 1.86 meters. In comparison, Euclidean has an accuracy rate of 99.65%, an average -the average error is 0.35% with the difference in average distance at the actual length of 2.42 meters. Therefore, the difference in distance between the two algorithms obtained is 1.27 meters. And the K-Means Algorithm can be the proper method for making decisions because the algorithm groups according to the farthest, medium, and closest distances

Copyright © 2022 Universitas Indraprasta PGRI.
All rights reserved.

Corresponding Author:

Mustofa Kamal Syarifudin,
faculty of informatic,
Universitas Nasional,
Jl. Sawo Manila, Pejaten Ps. Minggu Jakarta 12520
Email: 2018.mutofa.kamal.syarifudin@student.unas.ac.id

1. PENDAHULUAN

Pendidikan Nasional merupakan pendidikan yang diatur oleh undang-undang Republik Indonesia mulai dari pelaksanaan penerimaan, [1] pelajaran hingga kegiatan di sekolah. Peraturan dalam pendidikan nasional ini mengatur tentang penerimaan peserta didik baru yang akan difokuskan pada peraturan daerah tentang sistem penerimaan jalur zonasi, dimana jalur penerimaan ini diatur sedemikian rupa bagi peserta didik baru untuk memilih sekolah yang dekat dengan tempat tinggal peserta didik tersebut. [2]

Peraturan tentang zonasi tersebut dibuat untuk membagi kebutuhan secara seimbang untuk pendidikan disetiap sekolah pada daerah tersebut, dibutuhkan keakuratan dalam menentukan jarak antara tempat tinggal dengan sekolah yang akan menerima pendaftaran. Maka dari itu penulis membuat sebuah aplikasi yang diimplementasikan dengan menentukan titik koordinat menggunakan GPS di perangkat peserta didik yang diolah dengan algoritma euclidean distance dan membandingkan dengan algoritma haversine distance untuk menentukan keakuratan jarak antara dua titik dan juga divisualisasikan titik tersebut dengan API google maps. [3]

Pada penelitian ini penulis memfokuskan pada Sekolah Menengah Pertama di daerah kabupaten tangerang yaitu SMPN 1 Tigaraksa di Jalan Ki Mas Laeng no. 70 Kelurahan Tigaraksa Kecamatan Tigaraksa Kabupaten Tangerang dengan titik koordinat 6°15'48.5"S 106°28'04.2"E, Setelah mendapatkan data yang

diinginkan peneliti membuat klustering pada hasil tersebut dengan menggunakan metode K-Means agar dapat menentukan jarak radius terpendek dari sekolah SMPN 1 Tigaraksa ke tempat tinggal peserta didik [4]. Kebijakan zonasi ini harus disambut baik karena pemerintah ingin semua warga Indonesia bisa bersekolah dan memiliki pendidikan yang cukup, serta terjangkau dari tempat tinggal.

2. METODE

2.1. Penelitian Terkait

Pada tahun 2020, Yusuf dkk melakukan penelitian tentang perbandingan metode perhitungan jarak euclidean, haversine, dan manhattan dalam penentuan posisi karyawan yang dilakukan di Institut Teknologi Nasional Bandung yang menghasilkan perhitungan metode Euclidean distance dan Haversine distance memiliki rata-rata selisih jarak dengan perhitungan sebenarnya sebesar kurang dari 0,5 meter. Sedangkan manhattan memiliki rata-rata selisih 6,67 meter. [5]

M. Nishom pada tahun 2019 meneliti tentang perbandingan jarak antara metode euclidean, manhattan, dan Minkowski untuk mengetahui disparitas kebutuhan guru di kota Tegal, dengan hasil tingkat akurasi untuk euclidean distance sebesar 84,47%, lalu untuk manhattan distance sebesar 83,85%, dan untuk Minkowski sebesar 83,85% maka disimpulkan bahwa metode euclidean merupakan metode terbaik dari perbandingan ketiga metode tersebut. [6]

2.2. Penentuan Jarak

Jarak pada penelitian ini ditentukan dari 2 titik koordinat yang nanti akan diolah untuk mendapat hasil jarak yang diinginkan, maka dari itu untuk menentukan titik koordinat pada umumnya menggunakan *latitude* dan *longitude*. Kita bisa tahu dimana garis yang melintang diantara kutub utara dan selatan juga menghubungkan sisi timur dan barat bagian bumi itu bisa disebut dengan *latitude*, sedangkan *longitude* adalah garis yang menghubungkan utara dan selatan bumi juga bisa disebut garis meridian. [7]

Penentuan jarak ini penulis memanfaatkan dengan teknologi masa kini yaitu dengan GPS di dalam *smartphone* atau *gadget* yang mendukung sistem GPS, yang nanti mengirimkan sinyal ke satelit untuk menentukan titik koordinat.

2.3. Euclidean

Metrik Euclidean (dan besaran jarak) adalah yang sesuai dengan pengalaman sehari-hari dan persepsi. Yaitu, jenis metrik linier 1, 2, dan 3 Dimensi dunia dimana jarak antara dua titik dalam ruang sesuai dengan panjang garis lurus yang ditarik di antara keduanya. [7]

Berikut adalah persamaan dari *euclidean distance* :

$$d = \sqrt{(lat1 - lat2)^2 + (long1 - long2)^2} \times (1 \text{ derajat bumi}) \quad (1)$$

Keterangan :

- d = Jarak
- $lat1$ = Koordinat Latitude 1
- $lat2$ = Koordinat Latitude 2
- $Long1$ = Koordinat Longitude 1
- $Long2$ = Koordinat Longitude 2

2.4. Haversine

Teorema Haversine digunakan untuk menghitung panjang dua titik pada permukaan bumi berdasarkan pada garis lintang dan bujur. [3]

Berikut adalah persamaan dari *haversine distance* :

$$d = 2r \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta lat}{2} \right) + \cos(lat1) \cos(lat2) \sin^2 \left(\frac{\Delta long}{2} \right)} \right) \quad (2)$$

Keterangan :

- d = Jarak
- r = Jari – jari bumi
- Δlat = Besar perubahan latitude
- $\Delta long$ = Besar perubahan longitude

2.5. K-Means

K-Means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha membagi data ke dalam bentuk dua atau lebih kelompok. Metode ini membagi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berbeda karakteristik dikelompokkan ke dalam kelompok yang lain. [6], [8], [9]

Berikut langkah-langkah pada algoritma K-Means secara umum :

1. Menentukan banyaknya k kelompok.

2. Membagi data ke dalam k kelompok.
3. Menghitung pusat kelompok (sentroid) dari data yang ada di masing-masing kelompok dan dinyatakan dalam bentuk persamaan (3).

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j \quad (3)$$

dimana C adalah sentroid, M adalah banyak data, i adalah banyak kelompok

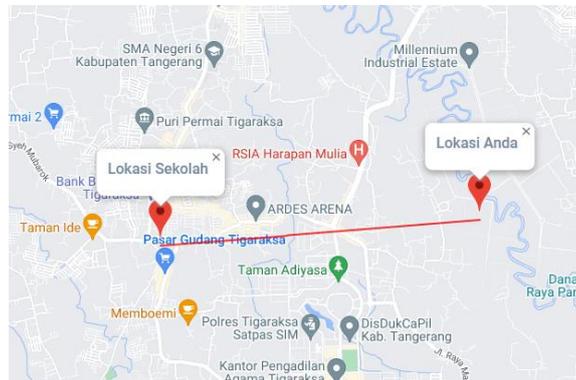
4. Masing-masing data dialokasikan ke sentroid terdekat. Menghitung jarak data ke setiap sentroid menggunakan jarak *Euclidean* dan dinyatakan dalam bentuk persamaan (4)

$$D(x_i, C_i) = \sqrt{\sum_{j=1}^q (x_{ij} - C_{ij})^2} \quad (4)$$

5. Kembali ke langkah 3 apabila masih ada data yang berpindah kelompok

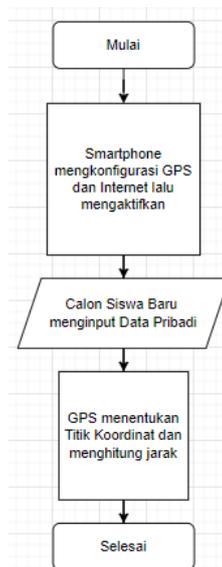
2.6. Google Map

Google Map adalah aplikasi peta gratis yang diberikan oleh perusahaan raksasa Google, yang dapat diakses melalui API Javascript agar pengguna dapat leluasa bereksperimen dengan produk Google Map yang telah disediakan. Informasi geografis dan perencanaan kota dapat dilihat secara rinci ketika peta terbuka. Layanan ini memiliki antarmuka interaktif. Posisi peta dapat digeser dan diperbesar sebagai diinginkan. Apalagi sekarang Google Maps sudah memiliki Street View. Ini adalah realitas virtual untuk jalan-jalan bumi ini. Selain itu, Google Maps juga menawarkan peta yang diputar di seluruh dunia, dan menawarkan perjalanan rute.[3]



Gambar 2. Tampilan alokasi jarak antara sekolah dengan lokasi rumah dengan Google Map

2.7. Diagram Alur Logika



Gambar 4. Alur Logika Pembuatan Sistem

Pada Gambar 4 diperlukannya logika untuk terkonsepnya pembuatan aplikasi yang ingin dibangun, di dalam alur logika tersebut terdapat perhitungan jarak yang akan ditentukan.

Berikut adalah titik pusat sekolah yang akan dijadikan acuan dalam menentukan jarak terpendek dari rumah ke sekolah, yang diambil dari posisi tengah pada sekolah dilihat dari google map.

Tabel 1. Titik pusat koordinat sekolah

| Latitude | Longitude |
|-------------------|-------------|
| -6.2636992 | 106.4678345 |

2.8. Tahap Perhitungan Jarak

Data yang telah tersimpan di dalam database akan diolah untuk mendapatkan jarak antar titik koordinat data calon siswa dengan titik koordinat pusat sekolah. Dilakukan dengan dua cara yaitu dengan algoritma *Euclidean* dan *Haversine*, lalu data yang telah dilakukan perhitungan jarak akan diklasifikasikan menggunakan metode K-Means sehingga sekolah dapat menarik hasil dari pendaftaran calon siswa jalur zonasi. Berikut contoh perhitungan jarak menggunakan setiap metode terhadap data yang ada pada tabel 2.

Tabel 2. Sampel Data Koordinat

| Koordinat | Latitude | Longitude |
|--------------------|------------|-------------|
| Sekolah | -6.2636992 | 106.4678345 |
| Calon Siswa | -6.2636500 | 106.4720899 |

Cara menghitung *euclidean distance* dari titik koordinat pada Tabel 2. Dari persamaan (1) bisa kita rumuskan sebagai berikut :

$$d = \sqrt{((-6.2636992) - (-6.2636500))^2 + (106.4678345 - 106.4720899)^2} \times (111322)$$

$$d = 473,75 \text{ Meter}$$

Lalu cara menghitung *haversine distance* dari titik koordinat pada Tabel 2. Dari persamaan (2) bisa kita rumuskan sebagai berikut

$$d = 6371111. \left(\sqrt{\sin^2\left(\frac{0,00004919}{2}\right) + \cos(-6.2636992)\cos(-6.2636500)\sin^2\left(\frac{0,004255}{2}\right)} \right)$$

$$d = 470,39 \text{ Meter}$$

3. HASIL DAN PEMBAHASAN

3.1. Penghitungan Ketepatan Jarak

Pada perhitungan ini akan dilihat tingkat ketepatan dengan membandingkan jarak yang dihasilkan oleh aplikasi dengan jarak yang diukur secara manual. Pengukuran dilakukan menggunakan pita ukur yang ada pada aplikasi Google Map dengan ketelitian 2 decimal, Berikut sampel perhitungan jarak.

Tabel 3. Sampel Data yang telah diolah dan dibandingkan

| Data | Lat | long | Dalam Hitungan Meter | | |
|-----------|-----------|------------|----------------------|-----------|----------|
| | | | Manual | Euclidean | Haversin |
| 1 | -6,267721 | 106,468283 | 448,56 | 450,50 | 449,96 |
| 2 | -6,266949 | 106,468729 | 373,73 | 375,24 | 374,66 |
| 3 | -6,277926 | 106,467936 | 1582,43 | 1583,80 | 1582,02 |
| 4 | -6,281165 | 06,467983 | 1995,58 | 1998,39 | 1996,14 |
| 5 | -6,274463 | 106,467933 | 1195,66 | 1198,30 | 1196,95 |
| 6 | -6,265141 | 106,467931 | 159,06 | 160,87 | 160,68 |
| 7 | -6,280633 | 106,467919 | 1882,93 | 1885,13 | 1883,01 |
| 8 | -6,264855 | 106,467901 | 128,12 | 128,88 | 128,73 |
| 9 | -6,280000 | 106,467999 | 1818,55 | 1814,73 | 1812,69 |
| 10 | -6,267956 | 106,467973 | 472,57 | 474,13 | 473,59 |

Dari hasil perbandingan data dari Tabel 3, dilakukan perhitungan persentasi Error untuk setiap hasil dengan persamaan berikut :

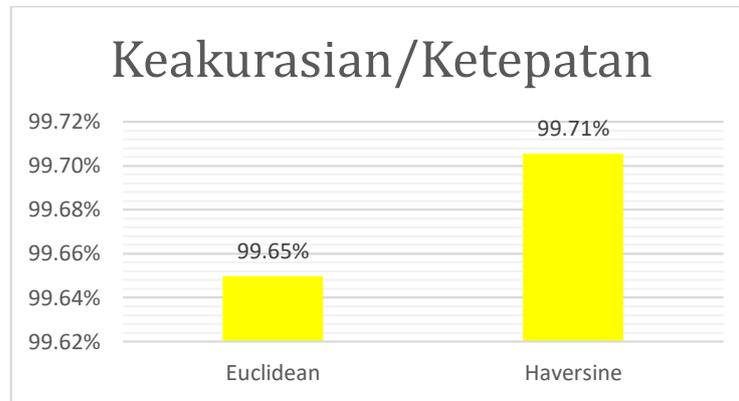
$$e = \frac{|a-b|}{a} \times 100\% \quad (5)$$

dimana a adalah nilai jarak sebenarnya dan b adalah nilai jarak hasil perhitungan aplikasi.

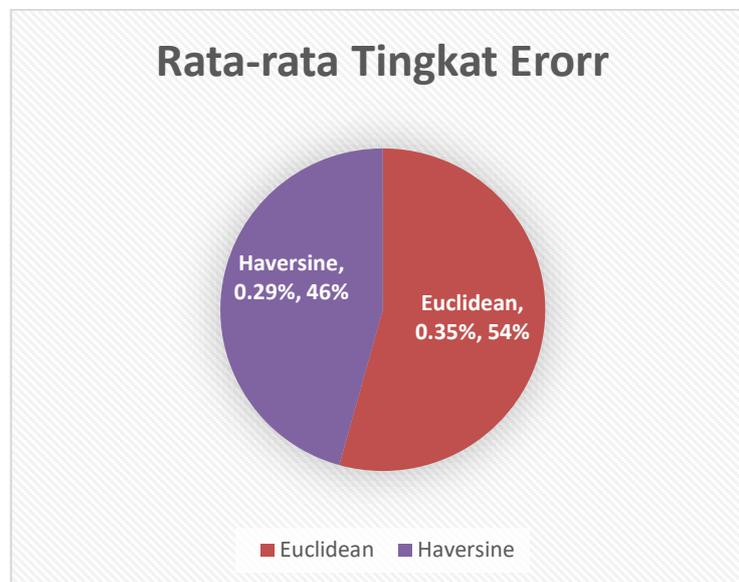
Lalu dilanjutkan dengan menghitung tingkat ketepatan dengan persamaan berikut:

$$\text{keakurasian} = 100\% - e \quad (6)$$

Nilai e adalah nilai error yang didapat.



Gambar 5. Grafik ketepatan Jarak dari kedua metode



Gambar 6. Grafik rata-rata tingkat erorr dari kedua metode

Dari hasil uji tingkat ketepatan pada setiap sampel data dapat dilihat bahwa dari kedua metode yang diuji metode *haversin* memiliki tingkat ketepatan dengan rata-rata 99,71% , rata-rata erorr 0,29% dengan rata-rata selisih jarak pada jarak sebenarnya 1,86 meter, sedangkan *euclidan* memiliki tingkat ketepatan 99,65% , rata-rata erorr 0,35% dengan rata-rata selisih jarak pada jarak sebenarnya 2,42 meter, adapun selisih jarak dari kedua metode yang didapat adalah 1,27 meter.

3.2. Klastering Data

Pada penentuan algoritma K-Means dihasilkan nilai centroid dari data yang diperoleh penulis menggunakan sempel data dari Tabel 3, adapun syarat centroid yang dipilih memiliki 3 kriteria yaitu Kriteria Dekat (C1), Sedang (C2), Jauh (C3). Penentuan titik ini ditentukan dari jarak terdekat, sedang, dan terjauh. Nilai titik tersebut dapat dilihat dari tabel 4 berikut :

Tabel 4. Centroid Data Awal

| Atribut | Nilai 1 | Nilai 2 |
|-----------------------|---------|---------|
| Klaster Dekat | 128,88 | 128,73 |
| Klaster Sedang | 375,24 | 374,66 |
| Klaster Jauh | 1198,29 | 1196,95 |

Dengan centroid yang telah didapat pada Tabel 4 maka dapat diklasterkan data yang telah didapat menjadi 3 klaster. Proses pengelompokan dengan menentukan jarak terdekat dari setiap data yang diolah. Berikut adalah pengolahan disetiap literasi

Tabel 5. Pengolahan Literasi-1

| Data | Cluster 2 | | Cluster 3 | | Cluster 1 | |
|--------------------|-----------|---------|-----------|---------|-----------|--------|
| | X | Y | X | Y | X | Y |
| 1 | 450,50 | 449,96 | 0,00 | 0,00 | 0,00 | 0,00 |
| 2 | 375,24 | 374,66 | 0,00 | 0,00 | 0,00 | 0,00 |
| 3 | 0,00 | 0,00 | 1583,80 | 1582,02 | 0,00 | 0,00 |
| 4 | 0,00 | 0,00 | 1998,39 | 1996,14 | 0,00 | 0,00 |
| 5 | 0,00 | 0,00 | 1198,30 | 1196,95 | 0,00 | 0,00 |
| 6 | 0,00 | 0,00 | 0,00 | 0,00 | 160,87 | 160,68 |
| 7 | 0,00 | 0,00 | 1885,13 | 1883,01 | 0,00 | 0,00 |
| 8 | 0,00 | 0,00 | 0,00 | 0,00 | 128,88 | 128,73 |
| 9 | 0,00 | 0,00 | 1814,73 | 1812,69 | 0,00 | 0,00 |
| 10 | 474,13 | 473,59 | 0,00 | 0,00 | 0,00 | 0,00 |
| Jumlah | 1299,87 | 1298,21 | 8480,34 | 8470,81 | 289,75 | 289,42 |
| Banyak data | 3 | 3 | 5 | 5 | 2 | 2 |
| Rata-rata | 433,29 | 432,74 | 1696,07 | 1694,16 | 144,87 | 144,71 |

Pada Literasi 1 masih bisa berubah pada banyak data disetiap kelompok dan literasi 1 belum bisa dipastikan pengelompokannya Maka dilanjut ke Literasi ke-2 untuk menentukan apakah data sudah tepat dan tidak akan berubah, jika data masih berubah pada Literasi ke-2 maka dilanjutkan ke Literasi-3 dan seterusnya.

Tabel 6. Pengolahan Literasi-2

| Data | Cluster 2 | | Cluster 3 | | Cluster 1 | |
|--------------------|-----------|---------|-----------|---------|-----------|--------|
| | X | Y | X | Y | X | Y |
| 1 | 450,50 | 449,96 | 0,00 | 0,00 | 0,00 | 0,00 |
| 2 | 375,24 | 374,66 | 0,00 | 0,00 | 0,00 | 0,00 |
| 3 | 0,00 | 0,00 | 1583,80 | 1582,02 | 0,00 | 0,00 |
| 4 | 0,00 | 0,00 | 1998,39 | 1996,14 | 0,00 | 0,00 |
| 5 | 0,00 | 0,00 | 1198,30 | 1196,95 | 0,00 | 0,00 |
| 6 | 0,00 | 0,00 | 0,00 | 0,00 | 160,87 | 160,68 |
| 7 | 0,00 | 0,00 | 1885,13 | 1883,01 | 0,00 | 0,00 |
| 8 | 0,00 | 0,00 | 0,00 | 0,00 | 128,88 | 128,73 |
| 9 | 0,00 | 0,00 | 1814,73 | 1812,69 | 0,00 | 0,00 |
| 10 | 474,13 | 473,59 | 0,00 | 0,00 | 0,00 | 0,00 |
| Jumlah | 1299,87 | 1298,21 | 8480,34 | 8470,81 | 289,75 | 289,42 |
| Banyak data | 3 | 3 | 5 | 5 | 2 | 2 |
| Rata-rata | 433,29 | 432,74 | 1696,07 | 1694,16 | 144,87 | 144,71 |

pada literasi ke-2 sudah tidak berubah maka klastering selesai, dapat dilihat bahwa dari sampel data yang didapat untuk kelastering jarak terdekat yaitu data 6 dan 8 untuk yang jarak sedang terletak pada data 1, 2 dan 10 dan untuk jarak jauh yaitu 3, 4, 5, 7, 9.

4. PENUTUP

Dari hasil pengujian kedua metode yang dibandingkan antara euclidean dan haversine memiliki tingkat ketepatan yang berbeda tipis atau bisa dikatakan hampir sama karena rata-rata dari metode haversine memiliki tingkat ketepatan dengan rata-rata 99,71% , rata-rata erorr 0,29% dengan rata-rata selisih jarak pada jarak sebenarnya 1,86 meter, sedangkan euclidean memiliki tingkat ketepatan 99,65% , rata-rata erorr 0,35% dengan rata-rata selisih jarak pada jarak sebenarnya 2,42 meter, adapun selisih jarak dari kedua metode yang didapat adalah 1,27 meter, dan penentuan klastering data yang didapat dengan contoh sampel dengan 3 centroid yang telah melewati 3 kali literasi maka didapat kelastering jarak terdekat yaitu data 6 dan 8 untuk yang jarak sedang terletak pada data 1, 2 dan 10 dan untuk jarak jauh yaitu 3, 4, 5, 7, 9.

Maka dapat disimpulkan bahwa dari kedua metode tersebut dapat diimplementasikan pada penentuan jarak zonasi sekolah. Serta menjadi acuan algoritma untuk pemerintah dalam membangun aplikasi yang sesuai dan akurat. Dari pihak sekolah bisa menggunakan algoritma K-Means agar menjadi tolak ukur untuk menentukan pengambilan keputusan diterima atau tidaknya pada kebijakan zonasi. Untuk selanjutnya penelitian ini dapat dikembangkan ke tahap selanjutnya.

DAFTAR PUSTAKA

- [1] R. Indonesia, “Undang-Undang Tentang Sistem Pendidikan Nasional,” 2003.
- [2] P. Suryanti, A. A. Musadad, and Isawati, “EFEKTIVITAS PELAKSANAAN SISTEM ZONASI DALAM PENERIMAAN PESERTA DIDIK BARU (PPDB) SERTA PENGARUHNYA TERHADAP UPAYA MANAJEMEN MUTU PENDIDIKAN BERDASARKAN ASAS KEADILAN DI SMA NEGERI KABUPATEN KLATEN,” *Jurnal Candi*, vol. 20, no. 1, pp. 111–126, 2020, Accessed: Dec. 07, 2022. [Online]. Available: <https://jurnal.uns.ac.id/candi/article/view/41335>
- [3] S. Hartanto, M. Furqon, A. P. U. Siahaan, and W. Fitriani, “Haversine Method in Looking for the Nearest Masjid,” *International Journal of Recent Trends in Engineering and Research*, vol. 3, no. 8, pp. 187–195, Aug. 2017, doi: 10.23883/ijrter.2017.3402.pd61h.
- [4] K. D. Pendidikan, “KEPUTUSAN KEPALA DINAS PENDIDIKAN KABUPATEN TANGERANG NOMOR : 422.1/Kep.1866-Disdik/2021,” 2021.
- [5] Y. Miftahuddin, S. Umaroh, and F. R. Karim, “PERBANDINGAN METODE PERHITUNGAN JARAK EUCLIDEAN, HAVERSINE, DAN MANHATTAN DALAM PENENTUAN POSISI KARYAWAN,” *Jurnal Tekno Insentif*, vol. 14, no. 2, pp. 69–77, Aug. 2020, doi: 10.36787/jti.v14i2.270.
- [6] M. Nishom, “Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 4, no. 1, pp. 20–24, Jan. 2019, doi: 10.30591/jpit.v4i1.1253.
- [7] C. A. Pamungkas, “APLIKASI PENGHITUNG JARAK KOORDINAT BERDASARKAN LATITUDE DAN LONGITUDE DENGAN METODE EUCLIDEAN DISTANCE DAN METODE HAVERSINE,” *Jurnal INFORMA Politeknik Indonusa Surakarta*, vol. 5, no. 2, pp. 8–13, 2019.
- [8] C. Astria, A. P. Windarto, A. Wanto, and E. Irawan, “Metode K-Means Pada Pengelompokan Wilayah Pendistribusian Listrik,” *Seminar Nasional Sains & Teknologi Informasi (SENSASI)*, pp. 306–312, 2019, [Online]. Available: <http://prosiding.seminar-id.com/index.php/sensasi/issue/archivePage|306>
- [9] Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 5, no. 1, pp. 17–24, Apr. 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.