

Implementasi K-Means untuk Sistem Informasi Penyebaran Penyakit Epidemik di Kota Malang

Implementation of K-Means For Information Systems for the Spread of Epidemic Diseases in Kota Malang

Dyas Irvan Masruri¹
Sugeng Widodo²
Febry Eka Purwiantono³

^{1,2,3} Teknik Informatika, STIKI Malang, Indonesia
¹ dyas.mail@gmail.com, ² sugeng@stiki.ac.id, ³ febry@stiki.ac.id

***Penulis Korespondensi:**
Febry Eka Purwiantono
febry@stiki.ac.id

Riwayat Artikel:

Diterima : 12 Juli 2021
Direview : 27 Juli 2021
Disetujui : 3 Desember 2021
Terbit : 3 Desember 2021

Abstrak

Epidemi adalah penyakit yang penularannya sangat cepat di daerah yang luas dan menimbulkan banyak korban sehingga diperlukan perawatan dan perhatian intensif di daerah yang terkena dampak epidemi secara cepat waktu dan tepat. Ada banyak penyakit menular yang terjadi di Kota Malang, salah satunya adalah Demam Berdarah *Dengue* (DBD). Dengan adanya informasi yang cepat dan jelas dapat membantu Dinas Kota Malang dalam mengambil keputusan terkait penanggulangan penyakit epidemi. Sehingga peneliti membuat sistem informasi geografis penyebaran penyakit epidemi menggunakan algoritma *k-means* berbasis web yang akan menyajikan informasi terkait penyakit epidemi di Kota Malang. Untuk menghitung algoritma *k-means* diperlukan kriteria sebagai variabel dan dalam penelitian ini kriteria yang diambil adalah jumlah penderita di setiap kelurahan selama setahun. Dari kriteria tersebut akan menghasilkan 3 kelompok yang dideklarasikan menggunakan C1 sebagai daerah aman, C2 sebagai daerah sedang, dan C3 sebagai daerah rawan. Hasil dari pengelompokan tersebut kemudian disajikan dalam bentuk peta Kota Malang dengan bantuan MapBox. Dalam peta tersebut kelompok C1 ditandai dengan warna hijau, kelompok C2 dengan warna kuning, dan kelompok C3 dengan warna merah.

Kata kunci: sistem informasi, epidemi, penyakit menular, *clustering*, *k-means*

Abstract

Epidemic is a disease that spreads very quickly in a large area and causes many victims who require intensive care and attention in areas affected by the epidemic quickly and appropriately. There are many infectious diseases that occur in Malang City, one of which is Dengue Hemorrhagic Fever (DHF). The availability of fast and clear information can help the Malang City Health Office in making decisions regarding the prevention of epidemic diseases. So the researchers created a geographic information system for the spread of epidemic diseases using a web-based k-means algorithm that will present information related to epidemic diseases in the city of Malang. To calculate the k-means algorithm, criteria are needed as variables and in this study the criteria taken are the number of patients in each kelurahan for a year. From these criteria, 3 groups will be declared using C1 as a safe area, C2 as a moderate area, and C3 as a vulnerable area. The results of the grouping are then presented in the form of a map of Malang City with the help of MapBox. In the map, group C1 is marked in green, group C2 is in yellow, and group C3 is in red.

Keywords: information system, epidemic, infectious diseases, *clustering*, *k-means*

1. Pendahuluan

Epidemi adalah keadaan dimana terdapat frekuensi penyakit melebihi frekuensi biasa, atau dalam waktu singkat terdapat penyakit yang berlebih [1]. Dengan kata lain, epidemi adalah wabah yang terjadi secara lebih cepat daripada yang diduga. Salah satu contoh penyakit epidemi adalah wabah DBD (Demam Berdarah Dengue) di Indonesia. DBD adalah penyakit yang disebabkan oleh gigitan nyamuk *Aedes Aegypti*, dengan ciri demam tinggi mendadak disertai pendarahan dan dapat menimbulkan kematian.

Berdasarkan data dari Kementerian Kesehatan Republik Indonesia kasus penderita DBD di Indonesia yang tercatat sejak awal tahun 2019 hingga 29 Januari 2019 13.683 kasus di seluruh Indonesia. Dari jumlah tersebut yang dilaporkan meninggal adalah 133 orang. Dan adanya kenaikan yang tinggi pada 3 Februari 2019, yaitu dilaporkan jumlah penderita DBD mencapai 16.692 kasus dengan 169 orang meninggal. Jawa Timur menjadi provinsi jumlah kasus terbanyak [2]. Sementara itu kasus demam berdarah yang terjadi di Kota Malang pada Januari 2019 sebanyak 52 orang. Dari penderita tersebut tidak ada korban jiwa. Jumlah kasus terbanyak terjadi pada tahun 2016, yaitu sebanyak 464 kasus demam berdarah. Sejak tahun 2019 Dinas Kesehatan (Dinkes) Kota Malang melakukan gerakan satu rumah satu jumatik guna menanggulangi kasus DBD. Selain itu masyarakat juga dihimbau untuk menjaga kebersihan lingkungan [3].

Informasi tentang penyebaran wabah penyakit di Kota Malang akan terlihat lebih jelas dengan menampilkan peta. Dengan informasi yang jelas maka dapat membantu Dinas Kesehatan Kota Malang dalam pengambilan keputusan terkait penanggulangan. Secara umum peta dalam bentuk web mampu mengolah data berbentuk spasial dan data atribut yang dapat memberikan informasi jarak akurat. Sebelum data diolah kedalam peta perlu adanya pengelompokan data menjadi 3 bagian yaitu sangat rawan, rawan, dan aman dengan menggunakan algoritma *k-means clustering*. *K-Means clustering* adalah algoritma yang menangani pengelompokan data dengan cara mempartisi data kedalam satu atau lebih *cluster*/kelompok tertentu.

Oleh karena itu Kota Malang membutuhkan sistem informasi pengolahan peta yang diharapkan mampu memberikan gambaran persebaran wabah penyakit di Kota Malang yang disajikan dalam bentuk peta. Dengan bantuan SIG diharapkan dapat memudahkan membaca sebuah informasi dengan lebih menarik. Dalam penelitian ini metode *K-Means* digunakan untuk mengelompokkan kelurahan-kelurahan di Kota Malang berdasarkan tingkat kejadian suatu wabah agar cepat dalam pencegahan dan penanggulangan wabah tersebut. Data yang diolah merupakan data inputan *realtime* selama kurun waktu setahun. Hasil informasi tersebut kemudian akan divisualisasikan dalam bentuk peta yang saat masih belum ada di Dinas Kesehatan Kota Malang.

2. Metode Penelitian

Metode penelitian dalam penelitian ini digunakan untuk menguraikan langkah-langkah yang dalam penelitian. Tujuannya adalah memperoleh data sebaran penyakit epidemi di Kota Malang dalam bentuk peta. Adapun langkah yang dilakukan meliputi poin-poin berikut

Analisis Masalah

Analisis masalah dilakukan untuk menentukan pokok permasalahan dari penelitian yang akan dilakukan. Dalam penelitian ini analisis masalah dilakukan untuk memanfaatkan algoritma *k-means* pada kasus epidemi di Kota Malang dan menyajikannya dalam bentuk peta.

Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan menggunakan 2 metode. Metode pertama yang digunakan adalah identifikasi masalah. Dalam hal ini, identifikasi masalah dilakukan pada dinas kesehatan kota Malang untuk mendapatkan data penyakit epidemi. Metode kedua yang digunakan untuk mengumpulkan data dalam penelitian ini adalah studi literatur. Studi literatur adalah pengumpulan data yang diperoleh dari buku, jurnal ilmiah, maupun sumber terpercaya lainnya seperti website Kementerian Kesehatan Indonesia. Adapun data yang diperoleh dari hasil observasi dan studi literasi disajikan pada tabel berikut.

Tabel 1 Data Penyebaran Penyakit DBD Tahun 2019

Kelurahan	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sep	Okt	Nov	Des
Klojen	0	0	7	3	2	0	1	0	0	0	0	0
Samaan	0	0	1	1	1	0	0	0	0	0	0	0
Rampal Celaket	1	0	1	1	1	0	1	0	0	0	0	0
Kidul Dalem	0	1	3	3	3	3	0	0	0	0	0	0
Sukoharjo	0	4	0	0	1	0	2	0	0	0	0	0
Kasin	0	1	1	0	4	1	0	0	0	0	0	0
Kauman	0	2	1	2	1	1	0	0	0	0	0	0
Oro-oro Dowo	3	4	1	0	3	3	0	0	0	0	0	0
Bareng	2	3	0	2	2	1	0	0	0	0	1	0
Gading Kasri	1	2	0	0	4	1	1	0	0	0	0	0
Penanggungan	0	2	0	0	5	0	0	0	0	0	0	0
Balearjosari	0	0	0	0	0	0	0	0	0	0	0	0
Arjosari	0	7	2	0	0	0	0	0	0	0	0	0
Polowijen	1	0	0	0	1	0	0	0	0	0	0	0
Purwodadi	0	0	0	0	0	0	0	0	0	0	0	0
Pandanwangi	4	5	1	2	6	0	0	0	0	0	0	0
Blimbing	0	2	5	3	3	1	1	0	0	0	0	0
Purwantoro	2	3	2	2	9	0	1	0	1	0	0	0
Bunulrejo	1	2	0	1	0	0	0	0	0	0	0	0
Kesatrian	1	0	0	0	0	0	0	0	0	0	0	0
Polehan	0	0	1	2	1	1	1	0	0	0	0	0
Jodipan	1	1	0	2	1	1	0	0	0	0	0	0
Kotalama	5	0	1	0	2	0	0	0	0	0	0	0
Mergosono	2	1	2	0	2	0	0	0	0	0	0	0
Bumiayu	2	0	3	1	2	0	0	0	0	0	0	0
Wonokoyo	2	0	0	0	0	0	0	0	0	0	0	0
Buring	2	0	0	0	0	0	1	0	0	0	0	0
Kedung Kandang	2	0	5	0	0	0	0	0	0	0	1	0
Lesanpuro	2	0	1	0	0	0	0	0	0	0	0	0
Sawojajar	1	0	0	3	2	2	0	0	0	0	0	0
Madyopuro	2	5	1	0	0	0	0	0	0	0	0	0
Cemorokandang	1	0	0	0	0	0	0	0	0	0	0	0
Arjowinangun	3	5	0	4	8	0	1	0	0	0	0	0
Tlogowaru	0	0	0	0	1	3	0	0	2	0	0	0
Tunggulwulung	0	0	0	1	2	0	0	0	0	0	1	0
Merjosari	3	1	3	1	1	1	1	0	0	0	1	0
Tlogomas	2	3	1	2	1	0	0	0	0	0	0	0
Dinoyo	2	6	1	0	2	0	0	0	0	0	0	0
Sumbersari	2	4	4	0	0	0	0	0	1	0	0	0
Ketawang Gede	1	2	1	2	0	1	0	0	0	0	0	0
Jatimulyo	0	5	0	4	2	1	0	0	0	0	0	0
Tunjungsekar	0	4	2	0	2	1	0	0	0	0	0	0
Mojolangu	1	3	1	0	0	2	0	1	0	0	0	0
Tulusrejo	0	1	0	2	1	2	0	0	0	0	0	0
Lowokwaru	1	7	8	4	5	3	2	0	0	0	2	0
Tasikmadu	1	1	0	0	0	0	0	0	0	0	0	0
Ciptomulyo	0	3	0	1	0	2	0	0	0	0	0	0
Gadang	0	1	0	2	0	0	0	0	0	0	0	0

Kebonsari	1	0	2	0	1	0	0	0	0	0	0	0
Bandungrejosari	0	6	4	6	6	8	1	2	0	0	0	0
Sukun	3	7	10	5	2	4	0	0	0	0	0	0
Tanjungrejo	1	1	0	5	4	3	1	1	1	1	0	0
Pisang Candi	2	2	1	1	2	1	1	0	0	0	0	0
Bandulan	1	2	0	3	1	4	2	0	0	0	0	0
Karang Besuki	3	2	1	1	6	4	0	1	0	0	0	0
Mulyorejo	0	0	2	2	2	1	0	0	0	0	0	0
Bakalan Krajan	0	0	0	0	1	0	0	0	0	0	0	0

Analisis Kebutuhan Fungsional

Analisis kebutuhan fungsional dilakukan untuk mengetahui kebutuhan fungsional yang akan diterapkan dalam sistem yang akan dibangun. Dalam hal ini analisis kebutuhan fungsional sistem yang menjadi pokok pembahasan adalah implementasi algoritma *k-means* untuk mengetahui peta persebaran penyakit epidemik di kota Malang.

3. Hasil dan Pembahasan

Implementasi Metode K-Means

Sebelum diimplementasikan ke dalam sistem yang akan dibangun, metode *k-means* yang telah dipilih terlebih dahulu dilakukan simulasi menggunakan data yang telah dikumpulkan pada tahap sebelumnya. Tujuan simulasi adalah untuk memudahkan dalam menyusun algoritma aritmetika yang nantinya akan diterapkan dalam aplikasi. Adapun langkah-langkah yang dikerjakan dalam implementasi metode *k-means* meliputi langkah berikut:

Menentukan jumlah cluster

Dalam sistem ini *cluster* akan dibagi menjadi 3 berdasarkan tingkat rawan daerah persebaran penyakit epidemik yaitu aman, sedang, dan rawan. Untuk sementara *cluster* tersebut diidentifikasi sebagai C1 untuk aman, C2 untuk sedang, dan C3 untuk rawan.

Menentukan centroid secara acak

Berdasarkan tabel 1 ditentukan *centroid* secara acak dari jumlah penderita yaitu C1 dengan menggunakan nilai terendah dari keseluruhan data karena C1 merupakan bagian dari daerah aman, C2 menggunakan nilai rata-rata dari keseluruhan data karena C2 merupakan bagian dari daerah sedang, dan C3 menggunakan nilai tertinggi dari keseluruhan data karena C3 merupakan bagian dari daerah rawan. Hasilnya yaitu untuk C1 adalah Kelurahan Balarjosari, C2 adalah Kelurahan Arjosari, dan C3 adalah Kelurahan Bandungrejosari.

Menentukan cluster terdekat

Untuk menghitung jarak dari objek ke *cluster* pada tahap ini menggunakan rumus *Euclidean distance*. Pada objek Kelurahan Klojen memiliki data penderita dari bulan Januari hingga Desember kemudian dihitung jaraknya dengan *cluster* pertama, kedua dan ketiga. Menurut **INPUT CITATION** adapun rumus untuk menghitung *cluster* terdekat adalah sebagai berikut :

$$j(v_1, v_2) = \sqrt{\sum_{k=1}^N (v_1(k) - v_2(k))^2}$$

Menentukan kembali pusat centroid baru

Centroid baru ditentukan dengan tujuan untuk melakukan perhitungan ulang terhadap jarak *cluster* terdekat dengan *centroid* yang baru. Adapun rumus untuk menentukan *centroid* pusat yang baru adalah dengan mencari rata-rata setiap *cluster*.

Langkah ini dilakukan berulang hingga diperoleh nilai *cluster* setiap data dengan *centroid* tidak mengalami perubahan. Dalam penelitian ini simulasi dilakukan sebanyak 3 kali perulangan dan diperoleh hasil akhir jarak *cluster* dengan *centroid* sebagai berikut.

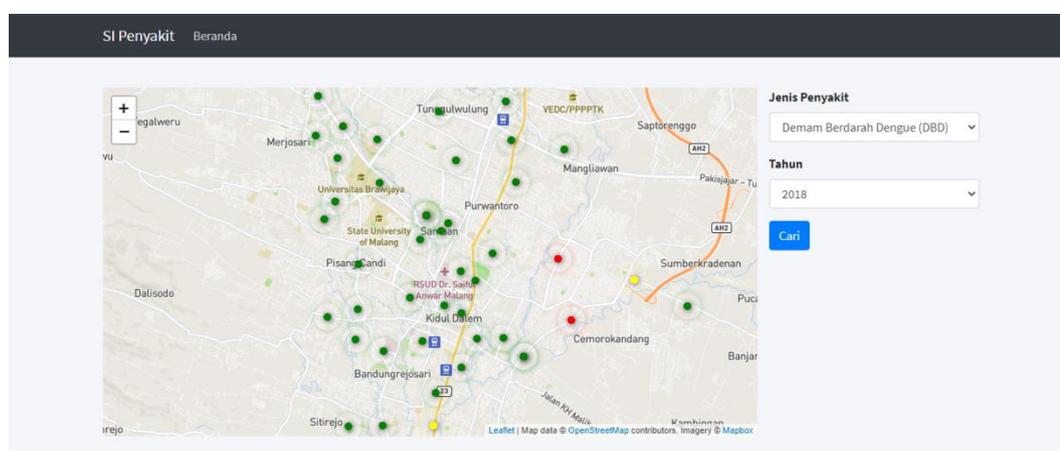
Tabel 2 Hasil akhir perhitungan *k-means cluster*

Kelurahan	Jarak dengan C1	Jarak dengan C2	Jarak dengan C3	Keterangan
Klojen	6,61	7,30	9,04	c1
Samaan	1,36	4,41	11,85	c1
Rampal Celaket	1,24	4,22	11,73	c1
Kidul Dalem	4,56	4,09	8,02	c2
Sukoharjo	4,21	3,79	11,21	c2
Kasin	3,55	3,19	10,82	c2
Kauman	2,13	2,99	10,08	c1
Oro-oro Dowo	5,20	3,06	9,09	c2
Bareng	3,42	2,13	9,98	c2
Gading Kasri	3,89	2,47	10,85	c2
Penanggungan	4,79	3,39	11,39	c2
Balearjosari	2,05	5,29	13,06	c1
Arjosari	6,59	5,55	10,05	c2
Polowijen	1,63	4,57	12,70	c1
Purwodadi	2,05	5,29	13,06	c1
Pandanwangi	7,55	4,41	9,42	c2
Blimbing	5,37	4,65	7,19	c2
Purwantoro	8,86	6,11	10,00	c2
Bunulrejo	1,97	3,90	11,72	c1
Kesatrian	1,78	5,13	13,00	c1
Polehan	1,97	4,26	11,11	c1
Jodipan	1,74	3,46	11,14	c1
Kotalama	4,36	5,45	12,46	c1
Mergosono	2,11	3,46	10,88	c1
Bumiayu	2,64	4,29	10,58	c1
Wonokoyo	2,05	5,17	13,01	c1
Buring	2,21	5,17	12,97	c1
Kedung Kandang	4,37	6,49	10,98	c1
Lesanpuro	1,73	5,06	12,48	c1
Sawojajar	3,19	4,09	10,89	c1
Madyopuro	4,62	4,21	10,68	c2
Cemorokandang	1,78	5,13	13,00	c1
Arjowinangun	9,36	6,10	9,98	c2
Tlogowaru	3,64	5,32	12,08	c1
Tunggulwulung	2,31	4,24	12,17	c1
Merjosari	3,10	4,28	9,87	c1
Tlogomas	2,87	2,71	10,05	c2
Dinoyo	5,65	3,64	9,92	c2
Sumbersari	4,79	4,98	9,50	c1
Ketawang Gede	2,01	3,57	10,38	c1
Jatimulyo	5,74	3,71	9,09	c2
Tunjungsekar	3,93	2,71	9,26	c2
Mojolangu	3,14	3,85	10,39	c1
Tulusrejo	2,45	3,78	10,89	c1
Lowokwaru	11,37	9,00	3,06	c3
Tasikmadu	1,69	4,56	12,52	c1
Ciptomulyo	3,24	3,83	10,71	c1

Gadang	2,14	4,51	11,93	c1
Kebonsari	1,52	4,55	11,66	c1
Bandungrejosari	12,36	9,82	5,32	c3
Sukun	12,44	10,90	4,28	c3
Tanjungrejo	6,24	5,04	9,63	c2
Pisang Candi	2,31	2,06	10,03	c2
Bandulan	4,83	4,53	9,68	c2
Karang Besuki	6,81	4,64	9,27	c2
Mulyorejo	2,34	3,94	10,34	c1
Bakalan Krajan	1,92	4,75	12,77	c1

Implementasi Program

Setelah melakukan simulasi perhitungan manual k-means dengan data hasil observasi, langkah selanjutnya adalah melakukan implementasi kepada sistem yang akan dibangun. Berikut ini hasil dari implementasi pada sistem yang dibangun.



Gambar 1 Hasil implementasi penyebaran penyakit

Pengujian

Setelah program selesai dibuat, maka terdapat proses uji coba pada sistem tersebut. Pengujian bertujuan untuk mengukur kualitas produk sebelum benar-benar digunakan khalayak umum. Berikut ini hasil pengujian pada sistem yang baru dibangun.

Tabel 3 Hasil Pengujian

Test Case	Hasil Harapan	Hasil Keluaran	Hasil Uji
Tekan tombol cari	Sistem melakukan tindakan	Sistem menampilkan peta sebaran penyakit berdasarkan jenis penyakit dan tahun	Sesuai

4. Penutup

Setelah dilakukan rancang bangun SIG untuk penyebaran penyakit epidemi di Kota Malang menggunakan algoritma *k-means*, maka dapat diambil kesimpulan sebagai berikut. SIG penyebaran penyakit epidemi yang menangani pendataan penderita hingga proses pengelompokan kelurahan telah berhasil dibangun. Implementasi sistem dibangun dengan Bahasa PHP, Python, Java, dan Javascript dengan bantuan Visual Studio Code. Sedangkan visualisasi SIG peta Kota Malang menggunakan Application Interface Programming (API) dari MapBox. Dari hasil perhitungan yang dihasilkan oleh sistem dan manual dapat disimpulkan bahwa menunjukkan hasil yang sama. Sistem yang telah

dibangun telah dilakukan pengujian dari segi web maupun android serta mendapat hasil yang sesuai dengan kriteria.

5. Referensi

- [1] F. Yanti, S. A. Lestari and M. G. Nangi, Dasar Epidemiologi, Yogyakarta: Grup Penerbitan CV BUDI UTAMA, 2019.
- [2] Kemenkes RI, "Kementrian Kesehatan Republik Indonesia," 04 Februari 2019. [Online]. Available: <https://www.kemkes.go.id/article/view/19020600004/kasus-dbd-terus-bertambah-anung-imbau-masyarakat-maksimalkan-psn.html>.
- [3] Jawa Pos, "JawaPos," 06 Februari 2019. [Online]. Available: <https://www.jawapos.com/jpg-today/06/02/2019/kasus-demam-berdarah-di-kota-malang-meningkat/>.
- [4] E. Y. Anggraeni and R. Irviani, Pengantar Sistem Informasi, Yogyakarta: CV. Andi Offset, 2017.
- [5] E. Irwansyah and M. Faisal, Advanced Clustering : Teori dan Aplikasi, Yogyakarta: Deepublish, 2015.