

PENGOLOMPOKAN DATA TINGKAT PARTISIPASI MASYARAKAT DALAM PENANGGULANAGAN HIV/AIDS MENGGUNAKAN K- MEANS BERBASIS MEDIAN DALAM PENENTUAN PUSAT KLASTER

Rizal Sulton ¹, Ahmad Supriadi ², Moh. Sukron ³
^{1,2,3} Universitas Nurul Jadid, Indonesia

Info Artikel	ABSTRAK
<p><u>Riwayat Artikel</u> Diterima: 20-03-2023 Disetujui: 29-04-2023</p> <p><u>Kata Kunci</u> <i>K-Means Berbasis Median; Clustering Tingkat Partisipasi, DBI</i></p> <p style="font-size: small; margin-top: 20px;">rizal.sulton@unuja.ac.id</p>	<p>Pada tahun 2016, di Indonesia kasus baru infeksi HIV meningkat 33,4% dibandingkan tahun sebelumnya. Terdapat 69,3% kasus baru infeksi HIV pada kelompok umur 25-49, sementara 63,3% penderita adalah laki-laki. Rasio HIV/AIDS antara laki-laki dan perempuan tercatat pada kisaran 2:1. Salah satu layanan penanggulangan HIV-AIDS dan IMS ialah konseling dan tes HIV (KT HIV). Berdasarkan berlimpahnya data konseling dan tes, informasi yang tersembunyi dapat diketahui dengan cara melakukan pengolahan terhadap data tersebut sehingga berguna bagi petugas kesehatan. Pengolahan data klien perlu dilakukan untuk mengetahui informasi penting berupa pengetahuan baru (knowledge discovery) yakni berupa pengelompokan data atau kluster data tingkat partisipasi masyarakat dalam penanggulangan HIV AIDS . Metode yang digunakan dalam pengolahan data di penelitian ini adalah K-Means dengan berbasis perhitungan median dalam penentuan pusat kluster, penggunaan median adalah untuk meningkatkan kinerja K-means dalam penentuan pusat kluster sehingga memperoleh hasil data pengelompokan lebih baik. Hasil penelitian menunjukkan bahwa proses kluster K-Means yang menggunakan penentuan pusat kluster berbasis median menghasilkan nilai evaluasi DBI 0,5857 sedangkan proses kluster K-Means tanpa penentuan pusat kluster berbasis median bernilai evaluasi DBI -0,647.</p>

1. PENDAHULUAN

Penyebab penyakit Acquired Immunodeficiency Syndrome atau yang sering disebut (AIDS) adalah Human Immunodeficiency Virus (HIV)[1]. Cara kerja virus tersebut dengan cara menyerang terhadap sel darah putih sehingga mengakibatkan rusaknya sistem imun(kekebalan) tubuh manusia. Fenomena gunung es ialah kata yang tepat untuk menggambarkan kasus HIV/AIDS didunia, ini disebabkan kuantitas jumlah orang yang dilaporkan jauh lebih sedikit dibandingkan dengan yang data sebenarnya. Hal ini tampak dari jumlah kasus AIDS yang dilaporkan, setiap tahunnya sangat meningkat secara signifikan. Di seluruh dunia, setiap hari diperkirakan sekitar 1400 anak di bawah usia 15 tahun meninggal dunia dan sekitar 2000 anak di bawah 15 tahun tertular HIV serta menginfeksi lebih dari 6000 orang berusia produktif [2].

Berdasarkan situs resmi pusat data Kementerian Kesehatan Republik Indonesia pada tahun 2010 2,1 juta orang hidup dengan HIV. Masuk pada tahun 2015 jumlah semakin meningkat

3,4juta, sekitar 36,7 juta di tahun tersebut orang hidup dengan HIV/AIDS. Akan tetapi berdasarkan data laporan WHO terjadi penurunan kematian, menurut hasil catatan WHO sejak awal AIDS ditemukan sampai pada akhir tahun 2015 terdapat 34 juta orang meningeal dunia. Pada tahun 2010 terdata 1,5 juta jiwa meninggal dunia dan pada tahun 2015 terdata 1,1 juta jiwa, menurun 4 juta jiwa pada tahun tersebut.

Melalui Kementerian Kesehatan Republik Indonesia, terkait dengan masalah HIV/AIDS pemerintah Indonesia meluncurkan program Layanan Komprehensif Berkesinambungan ialah sebuah kegiatan pelayanan pertama yang meliputi pencegahan penularan HIV/AIDS, pengobatan pasien HIV/AIDS, pemberian informasi HIV/AIDS ke masyarakat dan perawatan ODHA [3]. Pelayanan komprehensif dan berkesinambungan merupakan layanan HIV dan IMS yang lengkap. Pelayanan tersebut dilakukan mulai dari rumah atau komunitas, ke fasilitas layanan kesehatan seperti klinik, pukesmas dan rumah sakit dan sampai kembali kerumah atau komunitas sekaligus juga selama perjalanan infeksi HIV (semenjak sebelum terinfeksi sampai dengan stadium terminal). Kegiatan ini harus melibatkan seluruh pihak terkait, baik pemerintah, swasta, maupun masyarakat (kader, LSM, kelompok dampingan sebaya, ODHA, keluarga, PKK, tokoh adat, tokoh agama dan tokoh masyarakat serta organisasi/kelompok yang ada di masyarakat).

Salah satu layanan LKB HIV/AIDS adalah Konseling dan Tes HIV [3] sukarela. Pada prakteknya proses konseling dan tes akan didapatkan kumpulan data partisipasi masyarakat dalam penanggulangan HIV/AIDS. Berdasarkan berlimpahnya data konseling dan tes, tentunya terdapat informasi penting dan bermanfaat yang dapat diketahui dengan cara melakukan pengolahan terhadap data tersebut sehingga berguna bagi petugas kesehatan. Teknik pengolahan data hasil dari konseling dan tes tersebut perlu dilakukan untuk mengetahui informasi penting berupa pengetahuan baru (knowledge discovery)[4], contohnya informasi mengenai pengelompokan data tingkat partisipasi masyarakat dalam penanggulangan HIV/AIDS. Pengetahuan baru tersebut dapat membantu pihak pemangku kebijakan pada bidang kesehatan untuk mengetahui puskesmas daerah mana saja yang minim partisipasi masyarakat dalam penanggulangan HIV/AIDS, sehingga daerah tersebut mendapat perhatian khusus dan dicarikan solusi terbaik untuk mengatasi masalah tersebut.

Salah satu metode pengolahan data untuk mendapatkan informasi atau pengetahuan baru ialah metode *Clustering*. Klastering adalah proses pengelompokan sekumpulan data ke dalam beberapa klaster sehingga data yang terdapat pada satu klaster memiliki tingkat kemiripan yang maksimum dan data antar klaster memiliki kemiripan yang minimum[5]. Klastering mengacu pada pembagian data ke dalam kelompok objek yang sama. Setiap kelompok, terdiri dari elemen-elemen yang mirip satu sama lain dan berbeda dengan objek dalam kelompok lain. Klastering dalam proses pengolahan data berguna untuk menemukan pola pembagian di dalam sebuah kumpulan data yang berguna untuk proses pengkajian data, suatu bentuk pemodelan data yang bersumber pada ilmu matematika dan statistik. Dari perspektif pembelajaran mesin, klastering banyak digunakan untuk mengenali pola yang tersembunyi, model klastering merupakan pembelajaran tanpa pengawasan, dan sistem yang dihasilkan merupakan konsep data[6]. Kesamaan suatu objek biasanya didapat dari kedekatan nilai-nilai keunikan yang menjelaskan objek-objek data, sedangkan objek-objek data biasanya digambarkan sebagai sebuah titik dalam ruang multidimensi. Klastering dapat melakukan proses pengenalan daerah yang padat, pola-pola alokasi secara keseluruhan dan keterkaitan antara atribut-atribut data. Klastering berguna untuk beberapa analisis eksplorasi pola, pengelompokan, pengambilan keputusan, dan situasi pembelajaran mesin, termasuk data mining, pengambilan dokumen, segmentasi citra, dan klasifikasi pola.

K-means ialah salah satu prosedur metode klastering yang cukup sederhana dan populer untuk memisahkan kumpulan data ke dalam beberapa klaster k. Prosedur K-Means cukup mudah untuk diterapkan relatif cepat, mudah beradaptasi dan banyak dipergunakan [5]. Prinsip utama dari prosedur ini adalah menyusun data ke dalam klaster dengan mengelompokkan data berdasarkan kedekatan jarak antara data dari sekumpulan data yang ada dalam dataset. Prosedur

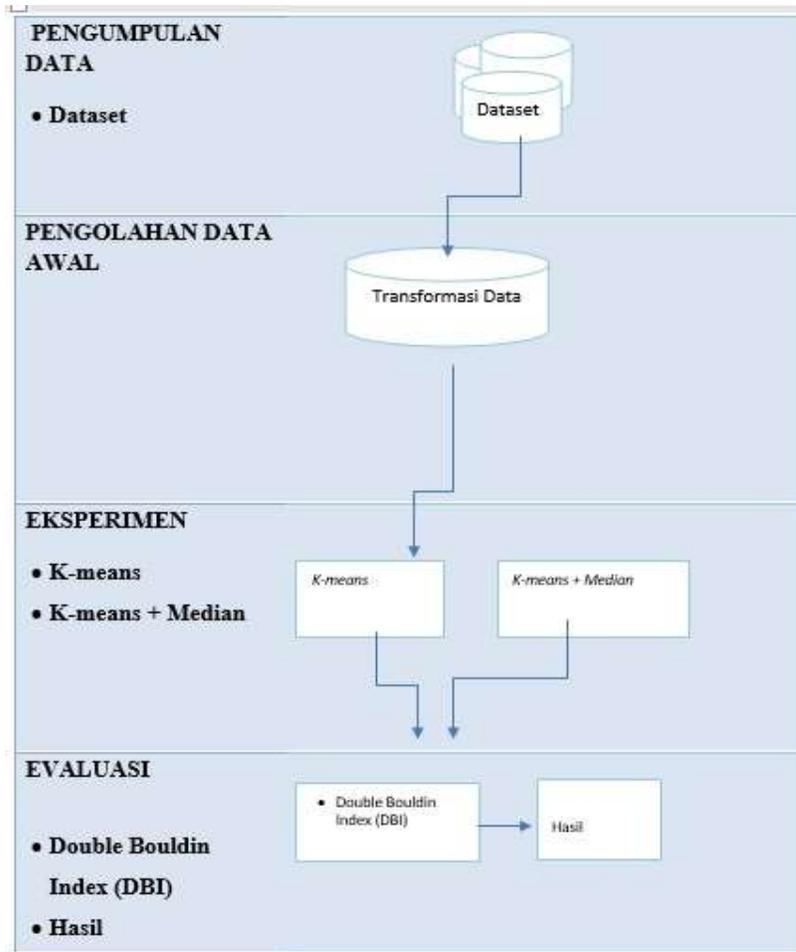
K-means dimulai dengan pembentukan pemisah kluster di awal, kemudian mengelompokkan data berdasarkan kedekatan jarak antar data secara berulang, pemisah kluster ini diperbaiki hingga tidak terjadi perubahan yang relevan pada pemisah kluster [8]. Prosedur K-means merupakan algoritma yang paling sering terkenal dan paling banyak digunakan untuk proses klastering data [8], [9]. Popularitasnya dapat dikaitkan dengan beberapa alasan. Pertama, secara konseptual sederhana dan mudah diimplementasikan. Hampir setiap software data mining mencakup pelaksanaan itu. Kedua, serbaguna yaitu hampir setiap aspek dari algoritma (inisialisasi, fungsi jarak, kriteria terminasi, dll) dapat dimodifikasi. Hal ini dibuktikan oleh banyak publikasi selama beberapa tahun terakhir yang memperbaiki metode k-means dalam berbagai cara, akan tetapi di sisi lain k-means memiliki beberapa kelemahan yang signifikan yaitu hasil dari klasterisasi sangat bergantung pada pemilihan pusat kluster awal, sehingga ketika pemilihan pusat kluster awal buruk maka dengan mudah terjebak dalam poor local minima [10],[11]. K-means sangat sensitif terhadap penentuan pusat kluster awal, karena akan mengakibatkan kluster akhir tidak optimal dan bahkan memungkinkan kluster kosong jika tidak ada poin yang dialokasikan kedalam kluster selama iterasi berlangsung [12]. Untuk alasan yang sama, K-means sangat sensitif terhadap pemilihan pusat kluster awal. Efek samping dari inisialisasi yang tidak tepat adalah memungkinkan adanya kluster kosong, konvergensi lambat, dan kesempatan yang lebih tinggi untuk terjebak dalam poor local minima [13] Oleh karena itu, cukup penting bagi K-means untuk memiliki pusat kluster awal yang baik.

Pada penelitian ini digunakan algoritma K-Means dengan memanfaatkan perhitungan nilai median dalam penentuan kluster untuk pengelompokkan tingkat partisipasi masyarakat dalam penanggulangan penyakit HIV/AIDS sehingga diharapkan mendapatkan hasil pengelompokan data terbaik.

2. METODE

Tabel 1 Kerangka Pemikiran

Problem Berdasarkan latar belakang permasalahan yang ada di dapat dirumuskan permasalahannya sebagai berikut : Apakah penentuan pusat kluster dengan nilai Median dapat mengoptimalkan Metode K-Means dalam klasterisasi data partisipasi masyarakat dalam penanggulangan HIV/AIDS sehingga dapat mengkluster dengan baik dan akurat.		
Sumber Data Data Layanan Tes dan Konseling HIV/AIDS dan IMS		
Penentuan Pusat Kluster dengan nilai median	METODE	Penentuan Pusat Kluster dengan nilai median
Evaluasi DBI (davies bouldin index)		



Gambar 1 Alur Tahapan Metode yang Diusulkan

3. HASIL DAN PEMBAHASAN

Menurut teori konsep DBI, DBI tersebut menginginkan nilai (non-negatif ≥ 0) sekecil mungkin untuk menilai baiknya kluster yang didapat. Dengan demikian berdasarkan hasil evaluasi DBI pada proses K-means menggunakan penentuan pusat kluster awal berbasis median lebih baik dari pada proses perhitungan K-Means tanpa menggunakan penentuan awal kluster berbasis median K-Means + Median

Tabel 2 Kerangka Pemikiran

R	1	2	3	Max(Rx,y)
1	0	0,595623	0,347949	0,595623
2	0,595623	0	0,565907	0,595623
3	0,347949	0,565907	0	0,565907

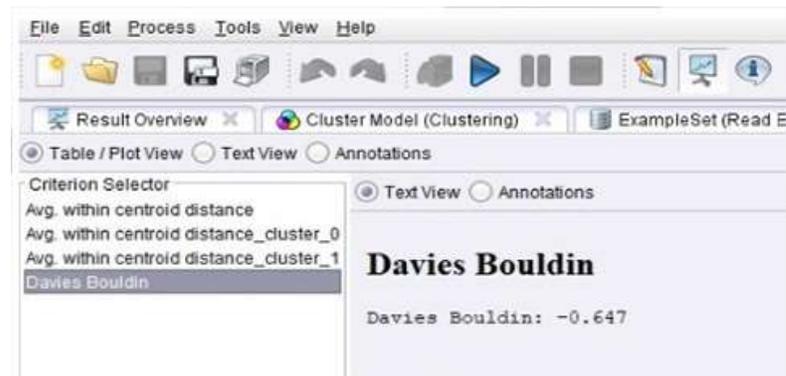
Tabel 1 adalah tahapan proses untuk mendapatkan nilai evaluasi DBI (Davies – Bouldin Index). Dengan k adalah nilai jumlah kluster dan Max-R adalah nilai tertinggi dari $R_{1,2}, R_{1,3}$ dan $R_{2,3}$. Untuk hasil perhitungan DBI (Davies – Bouldin Index) sebagai berikut :

$$DBI = \frac{1}{3} (0,595 + 0,595 + 0,566)$$

$$DBI = 0,5857$$

Gambar 2 Hasil Perhitungan DBI (Davies – Bouldin Index)

Berikut adalah hasil dari perhitungan menggunakan K-Means tanpa median



Gambar 3 implementasi Davies – Bouldin Index

4. KESIMPULAN

Dalam penelitian ini dilakukan pengujian model dengan menggunakan metode k-means dengan pemilihan pusat kluster awal berbasis median untuk mengurangi ketergantungan terhadap pemilihan pusat kluster awal. Hasil eksperimen metode K-Means dengan pemilihan pusat kluster awal berbasis median dari hasil eksperimen diatas dapat menghasilkan kualitas hasil kluster yang lebih baik, karena dengan pemilihan pusat kluster awal berbasis median tidak terjadi kluster kosong karena pemilihan pusat kluster awal yang tidak tepat, walaupun dari segi kecepatan proses masih kalah dari K-Means standart kluster yang dihasilkan yaitu pada kluster 1 didapat 108 kelompok pada kluster 2 didapat 422 kelompok dan kluster 3 didapat 490 sedangkan nilai DBI didapat 0,5857.

5. DAFTAR PUSTAKA

- [1] Lenny Octavianty, “Pengetahuan, Sikap Dan Pencegahan Hiv/Aids Pada Ibu Rumah Tangga,” Jurnal Kesehatan Masyarakat, 2015
- [2] Purwaningsih S. 2008. “Perkembangan HIV dan AIDS di Indonesia”. Jurnal Kependudukan Indonesia; 3 (2): 11 - 16
- [3] Desy Mulyasari “Layanan Komprehensif Berkesinambungan Terhadap Anak Dengan Hiv/Aids Di Kota Surakarta,” Jurnal Sosiologi DILEMA, Vol. 31, No. 1, 2016.
- [4] Rima Dias Ramadhani, “Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro”, Jurusan Sistem Informasi, Fakultas Ilmu Komputer Universitas Dian Nuswantoro
- [5] E. H. Witten, Ian H. Frank, "Data Mining Practical Machine Learning Tools and

- Techniques Third Edition". 2011.
- [6] P. Berkhin, "Survey of Clustering Data Mining Techniques," pp. 1–56.
- [7] E. H. Witten, Ian H. Frank, "Data Mining Practical Machine Learning Tools and Techniques Third Edition". 2011.
- [8] X. Wu and V. KUMar, The Top Ten Algorithms in Data Mining. 2009.
- [9] P. J. Flynn, "Data Clustering : A Review," vol. 31, no. 3, 2000.
- [10] J. A. Lozano and P. Larra, "An empirical comparison of four initialization methods for the K - Means algorithm," vol. 20, 1999. [8] M. E. Celebi, H. A. Kingravi, and P. A. Vela, "Expert Systems with Applications A.
- [11] M. E. Celebi, H. A. Kingravi, and P. A. Vela, "Expert Systems with Applications A.
- [12] S. Deelers and S. Auwatanamongkol, "Enhancing K-Means Algorithm with Initial Cluster Centers Derived from Data Partitioning along the Data Axis with the Highest Variance," vol. 26, no. December, pp. 323–328, 2007
- [13] M. E. Celebi, "A Comparative Study of Efficient Initialization Methods for the K-Means Clustering Algorithm," vol. 2, no. 1, 2012.
- [14] Soumi Ghosh, Sanjay Kumar Dubey," Comparative Analysis of K-Means and Fuzzy CMeans Algorithms", Vol. 4, No.4, 2013.
- [15] Navjot Kaur, Jaspreet Kaur Sahiwal, Navneet Kaur,"Efficient K-Means Clustering Algorithm Using Ranking Method In Data Mining", ISSN: 2278 – 1323,2012.
- [16] Hendi Shopian, "Pemilihan Pusat Klaster Awal Pada Metode K-Means Berbasis Median", ISSN Cetak: 2087-4286, ISSN On Line: 2580-6017.
- [17] Silvi Agustina dkk, "Clustering Kualitas Beras Berdasarkan Ciri Fisik Menggunakan Metode K-Means ", Program Studi Ilmu Komputer, Program Teknologi Informasi dan Ilmu Komputer Universitas Brawijaya Malang.
- [18] Agus Perdana Windarto," Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method", International Journal of Artificial Intelligence Research vol.No 2, 2017.
- [19] Mardalius," Implementasi Algoritma K-Means Clustering Untuk Menentukan Kelas Kelompok Bimbingan Belajar Tambahan", Proceeding SEMILOKA ROYAL 2017 "Teknologi Mobile",2017.
- [20] Eko Prasetyo, "Data Mining Mengolah Data Menjadi Informasi menggunakan matlab",Yogyakarta, Penerbit Andi,2014.
- [21] Wu, X. and Kumar, "The Top ten Algorithms in Data Mining". London,CRC Press taylor & Francis Group, 2009.
- [22] J. Isotalo, "Basics of Statistics," pp. 0–82.