PENGELOMPOKAN PASIEN GEJALA DEMAM BERDARAH DENGUE MENGGUNAKAN K-MEDOIDS CLUSTERING (Studi Kasus Rumah Sakit Citra Husada Jember)

Mila Yulia Saputri¹⁾, Agung Nilogiri²⁾, Deni Arifianto³⁾ Program Studi Teknik Informatika, Fakultas Teknik Universitas Muhammadiyah Jember

Jl. Karimata No. 49 Jember Kode Pos 68161

ABSTRAK

Demam Berdarah Dengue (DBD) merupakan penyakit yang disebabkan oleh virus dengue yang ditularkan melalui gigitan nyamuk *aedes aegypti. Clustering* adalah metode *data mining* yang dapat digunakan untuk pengelompkan pasien gejala DB D. Metode *clusiering* yang digunakan adalah metode *k*- medord.s berdasarkan model data hematokrit, leukosit, hemoglobin, tinmbosit dari pasien gejala DB D. Penelitian ini bertujuan untuk mencari jumlah cluster optimal berdarasrkan parameter hematokrit, leukosit, hemriglobin, tmmbosit menggunakan metrrde K-Medoid dan Djs Briuldin Index serta mencari jumlah masing-masing anggota cluster yang dihasilkan. Berdasarkan nilai Davies Bouldin fndex yang paling optimal adalah 6 cluster. Dengan nilai Davies bouldin Index sebesar 0,625.

Kata kunci: Demam Berdarah Dengue , Data Mining, Hematokrit, Leukosit, Hemoglobin, Tombosit, K-Medoid clustering

I. PENDAHULUAN

Penyakit Demam Berdarah Dengue (DBD) masih merupakan salah satu masalah masyarakat kesehatan yang utama Indonesia. DBD banyak ditemukan di daerah tropis dan sub-tropis termasuk di Indonesia, (Kemenkes RI, 2016) mencatat di tahun 2015 pada bulan oktober ada 3.219 kasus DBD kematian mencapai sementara November ada 2.921 kasus dengan 37 angka kematian, dan Desember 1.104 kasus dengan 31 kematian. Dibandingkan dengan tahun 2014 pada Oktober tercatat 8.149 kasus dengan 81 kematian, November 7.877 kasus dengan 66 kematian, dan Desember 7.856 kasus dengan 50 kematian. Kasus demam berdarah dengue (DBD) dikabupaten jember, jawa timur, sejak januari hingga pertengahan februari 2019 mencapai 148 kasus, namun tidak ada pasien meninggal dunia. Jumlah kasus DBD pada januari 2019 tercatat sebanyak 90 kasus dan februari hingga pertengahan tercatat sebanyak 50 kasus, "kata kepala dinas kesehatan (Dinkes) Jember, Siti Nurul Qomariyah di Jember, Senin. Menurutnya, kasus DBD terbanyak berada di Desa Kemuningsari Kidul.

Kecamatan Jenggawah, dan beberapa kelurahan di kecamatan Sumbersari, dan Kecamatan Kaliwates sehingga pihaknya terus gencar melakukan sosialiasi untuk pemberantasan sarang nyamuk (PSN). Jumlah kasus DBD memang mengalami peningkatan di kabupaten Jember dibandingkan periode yang sama tahun lalu, namun tidak ada penetapan kejadian luar biasa (KLB) terkait DBD di Jember (Dinkes, 2019).

Yayasan Citra Husada mendirikan rumah sakit Citra Husada pada bulan April tahun 2009 sebagai salah satu rumah sakit swasta di Jember. Rumah sakit Citra Husada diharapkan mampu menjawab keinginan masyarakat, khususnya masyarakat Jember dan sekitarnya terhadap pelayanan kesehatan yang bermutu, baik dari segi pelayanan kesehatan, sumber daya kesehatan yang profesional, sarana dan prasarana yang memenuhi standart yang ditentukan serta lokasi hunian rumah sakit yang nyaman, tenang dan keamanan terjamin. Rumah sakit citra husada jember juga menangani berbagai penyakit termasuk penyakit demam berdarah. Namun penanganan pasien dilakukan secara individual dan belum ada pengelompokan berdasarkan parameter gejala. Pentingnya dilakukan Cluster karena beda gejala juga akan beda penanganannya.

Menurut Han dan Kamber (2011), Clustering adalah proses pengelompokan kumpulan data menjadi beberapa kelompok sehingga objek didalam satu kelompok memiliki banyak kesamaan dan memiliki banyak perbedaan dengan objek dikelompok lain.perbedaan dan persamaannya biasanya berdasarkan nilai atribut dari objek tersebut dan dapat juga berupa perhitungan jarak. K-Medoids atau partitioning Arround Medoids (PAM) adalah algoritma clustering vang mirip dengan K-Means. K-Means menggunakan nilai rata-rata (mean) sebagai pusat cluster, sedangkan K-Medoid atau PAM mampu mengatasi menggunakan objek sebagai perwakilan (medoid) sebagai pusat cluster untuk setiap (Kaur,dkk.,2014). Algoritma K-Medoids memiliki kelebihan untuk mengatasi kelemahan pada algoritma K-Means yang sensitive terhadap noise dan outlier, dimana objek dengan nilai yang besar memungkinkan menvimpang pada distribusi data. Kelebihan lainnya yaitu hasil proses clustering tidak bergantung pada urutan masuk dataset (Furqon,dkk.,2015)

II. Data Mining

Datamining adalah sebuah proses menggunakan teknik statistik. yang matematika, kecerdasan buatan, machine mengekstraksi learning untuk mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban dkk. 2005). Data mining adalah rangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Pramudiono 2006). Terdapat beberapa istilah lain yang memiliki makna sama dengan data mining, vaitu knowledge discorvery in database (KDD), ekstraksi pengetahuan (knowledge extraction), Analisa data/pola (data/ pattern kecerdasan bisnis (business alaysis), intelegence) dan data archaeology dan data dredging (Larose, 2005).

Data mining mempunyai fungsi yang penting untuk membantu mendapatkan

informasi yang berguna serta meningkatkan pengetahuan bagi pengguna. Pada dasarnya, data mining mempunyai empat fungsi dasar diantaranya (Syarif et al.,2018):

- a) Fungsi Prediksi (prediction). Proses untuk menemukan pola dari data dengan menggunakan beberapa variabel untuk memprediksikan variabel lain yang tidak diketahui jenis atau nilainya.
- b) Fungsi Deskripsi (description). Proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data.
- Fungsi Klasifikasi (classification). Klasifikasi merupakan suatu proses untuk menemukan model atau fungsi untuk menggambarkan class atau konsep dari suatu Proses vang digunakan data. untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan.
- d) Fungsi Asosiasi (assosiation). Proses ini digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut dari sekumpulan data.

Tujuan data mining dapat dilihat dari dua sudut pandang, yaitu sudut pandang komersial dan sudut pandang keilmuan. Dari sudut pandang komersial, data mining dapat digunakan dalam menangani meledaknya volume data. Bagaimana menyimpannya, mengestraknya serta memanfaatkannya. Berbagai teknik komputasi dapat digunakan menghasilkan informasi yang dibutuhkan. Informasi yang dihasilkan asset untuk meningkatkan daya saing suatu institusi. Data mining tidak hanya digunakan menumpuknya menangani persoalan data/informasi bagaimana dan menggunakannya tanpa kehilangan informasi yang penting. Data mining juga diperlukan untuk menyelesaikan permasalahan atau menjawab kebutuhan bisnis itu sendiri, misalnya:

- a. Mengetahui hilangnya pelanggan karena pesaing.
- b. Mengetahui item produk atau konsumen yang memiliki kesamaan karakteristik.
- c. Mengidentifikasi produk-produk yang terjual bersamaan dengan produk lain.
- d. Memprediksi tingkat penjualan.

- e. Menilai tingkat resiko dalam menentukan jumlah produksi suatu item.
- f. Memprediksi perilaku bisnis dimasa yang akan datang.

III. Clustering

Clustering atau pengklasteran adalah suatu teknik data mining yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok sehingga data dalam satu kelompok atau cluster memiliki tingkat kemiripan maksimum, dan data antara cluster memiliki kemiripan yang minimum (Tan, 2006).

Metode yang terdapat pada teknik terdapat dua metode dalam cluster pengelompokkannya yaitu hierarchical clustering dan non-hierarchical clustering. Suatu metode pengelompokan data yang cara kerjanya dengan mengelompokan dua data atau lebih yang memiliki kesamaan atau kemiripan, kemudian proses dilanjutkan ke objek lain yang mimiliki kedekatan kedua, proses ini terus berlangsung hingga cluster membentuk semacam tree, dimana ada tingkatan yang jelas antar objek dari yang paling mirip hingga yang paling tidak mirip, merupakan pengelompokan data dengan cara Hierarchical clustering. Namun secara logika semua objek pada akhirnya hanya akan membentuk sebuah cluster (Santoso, 2010).

Sedangkan non-hierarchical clustering dimulai dengan menentukan jumlah cluster yang diinginkan (dua, tiga, atau lebih cluster), setelah jumlah yang cluster ditentukan maka proses cluster dimulai tanpa mengikuti proses hirarki, metode ini juga sering disebut sebagai metode K-Means Clustering (Santoso, 2010).

K-medoid atau Partitioning Around Method (PAM) adalah metode cluster non hirarki yang merupakan varian dari metode k-means. K-medoids hadir untuk mengatasi kelemahan K-Meaans yang sensitif terhadap outlier karena suatu objek dengan suatu nilai yang besar mungkin secara substansial menyimpang dari distribusi data (jiawei & kamber, 2006)

IV. Metode k-Medoids

Algoritma *K-Medoids* menggunakan metode partisi clustering untuk mengelompokkan sekumpulan objek menjadi sejumlah kcluster. Algoritma ini menggunakan objek pada kumpulan objek untuk mewakili sebuah cluster. Objek yang terpilih untuk mewakili sebuah cluster disebut dengan medoids. Cluster dibangun dengan menghitung kedekatan yang dimiliki antara medoid dengan objek non-medoid (Setyawati, 2017).

Menurut (Han and Kamber, 2006) Langkah-langkah algoritma *K-Medoids Clustering* adalah sebagai berikut:

- 1. Secara acak pilih *k* objek pada sekumpulan *n* objek sebagai *medoid*.
- 2. Secara acak pilih orandom : sebuah objek *non-medoid*.
- 3. Tempatkan objek *non-medoid* ke dalam *cluster* yang paling dekat dengan *medoid*.
- 4. Menghitung total *Cost*, S, dari pertukaran *medoid*Oj denganOrandom.
- 5. Jika S<0 maka tukar OjdenganOrandom untuk membentuk sekumpulan *k* objek baru sebagai *medoid*.
- 6. Ulangi langkah 2 hingga langkah 5 hingga tidak ada perubahan.

Nilai total, dan S dinyatakan dengan persamaan (Setyawati, 2017):

$$Total\ Cost = \sum_{k=1}^{n} (x_k - y_k)^2$$

.... (2.1)

Di mana:

n = jumlah suatu data.

k = indeks data.

 x_n = nilai atribut ke-k dari x.

yk = nilai atribut ke-k dari y

Nilai S dinyatakan dengan persamaan (Setyawati, 2017):

S = Totalcostbaru - Totalcostlama

Di mana:

S = Selisih.

Total *cost* baru = jumlah *cost non-medoid*.

Total *cost* lama = jumlah *cost medoid*.

V. Davies Bouldin Index (DBI)

David L.Davies dan Donald W. Bouldin memperkenalkan sebuah metode yang diberi nama dengan nama mereka berdua, yaitu Davies Bouldin Index (DBI) yang digunakan untuk mengevaluasi cluster (Narwin, et al. 2017). Evaluasi menggunakan Davies Bouldin index ini memiliki skema evaluasi internal cluster, dimana baik atau tidaknya hasil cluster dilihat dari kuantitas dan kedekatan antar data hasil cluster (Bates & kalita 2016).

Davies-Bouldin *Index* merupakan salah satu metode yang digunakan untukmengukur validitas cluster pada suatu metode pengelompokan, kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap titik pusat cluster dari cluster yang diikuti. Sedangkan separasi didasarkan pada jarak antar titik pusat cluster terhadap clusternya. Pengukuran dengan Davies Bouldin Index ini memaksimalkan jarak inter-cluster antara cluster Ci dan Ci, dan pada waktu yang sama mencoba untuk meminimalkan jarak antar titik dalam sebuah cluster. Jika jarak intermaksimal, berarti kesamaan cluster karakteristik antar masing-masing cluster sehingga perbedaan antar-cluster sedikit terlihat lebih jelas. Jika jarak intra-cluster minimal berarti masing-masing objek dalam cluster tersebut memiliki tingkat kesamaan karakteristik yang tinggi (Wani & Riyaz 2017). Oleh karena itu, nilai DBI yang lebih rendah berarti pengelompokan yang lebih baik.Tahapan dari perhitungan DaviesBouldin Index adalah sebagai berikut:

1. Sum of Square Within-cluster (SSW)

Untuk mengetahui kohesi dalam sebuah cluster ke-i adalah dengan menghitung nilai dari *Sum of Square Within-cluster* (SSW). Kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap titik pusat cluster dari sebuah cluster yang diikuti. Persamaan yang digunakan untuk memperoleh nilai *Sum of Square Withincluster* adalah sebagai berikut.

$$SSW_i = \sum_{j=i}^{mi} d(xj, c_i)$$

Dimana:
 X_{ji} =pusat atribut
 C_i = pusat Cluster

2. *Sum of Square Between-cluster* (SSB)

Perhitungan Sum of Square Betweencluster (SSB) bertujuan untuk mengetahui separasi antar *cluster* . Persamaan yang digunakan untuk menghitung nilai *Sum* of Square Between cluster adalah sebagai berikut.

$$SSB_{i, j} = (C_i, C_j)$$

Dimana:
 $C_i = \text{cluster ke-1}$
 $C_i = \text{cluster ke-2}$

3. Davies Bouldin Index

Nilai rasio yang diperoleh dari *Ratio* digunakan untuk mencari nilai *Davies-Bouldin Index* (DBI) dengan menggunakan persamaan berikut:

$$DB = \frac{1}{n} \sum_{i=1, i \neq j}^{n} \max(\frac{SSW_i + SSW_j}{SSB_i, j})$$
.....(2.3)

Di mana:

DB = Devies Bouldin

n = jumlah*cluster*.

 SSW_i = rata-rata jarak dari data dengan titik pusat data *cluster*ke-i.

 SSW_j = rata-rata jarak dari data dengan titik pusat data *cluster*ke-j.

 c_i = titik pusat data *cluster*i.

 c_i = titik pusat data *cluster*j.

 $SSB_{i,j}$ = jarak antara *centroid* atau *medoid* dari c_i dan c_j .

Menurut penelitian Meylindra Arini P., Rully A., Irmasari Hafidz, DBI paling rendah menunjukan cluster optimal .

VI. RapidMiner

RapidMiner merupakan sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner adalah perangkat lunak atau software yang bersifat terbuka (open source) yang berdiri sendiri untuk analisis data

sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. Dengan menggunakan berbagai teknik deskriptif prediksi dan RapidMiner memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik.

RapidMiner sebelumnya dikenal dengan nama YALE (Yet Another Learning Environment), yang awalnya dikembangkan oleh RalfKlinkenberg, Ingo Mierswa, dan Simon Fischer di Artificial Intelligence Unit dari University of Dortmund pada tahun 2001. RapidMiner didistribusikan di bawah lisensi AGPL (GNU Affero General Public License) versi 3. Hingga saat ini banyak ribuan aplikasi dikembangkan menggunakan yang RapidMiner di lebih 40 negara. RapidMiner menempati peringkat pertama pada polling oleh KDnuggets sebagai software data mining, sebuah portal data mining pada 2010-2011. Sehingga sebagai software open source untuk data mining tidak perlu diragukan lagi karena software ini sudah terkemuka di dunia.

RapidMiner dapat merancang sebuah pipeline analitis dengan menyediakan GUI (Graphic User Interface) yang menghasilkan file XML (Extensible Markup Language) yang mendefinisikan proses analitis keinginan pengguna untuk diterapkan ke data. RapidMiner kemudian membaca file ini untuk menjalankan analis secara otomatis (Aprilla dkk., 2013:8)

VII.IMPLEMENTASI PEMBAHASAN

DAN

Mencari Jumlah Cluster Terbaik

Pada penelitian ini jumlah cluster terbaik yang ditentukan berdasarkan perhitungan RapidMiner yaitu ada 6 cluster. 1. Bahwasannya tabel 4.2 adalah

1. Bahwasannya tabel 4.2 adalah perhitungan atau hasil dari 2 sampai 10 cluster

Tabel 4. 2 Perhitungan atau hasil dari 2 sampai 10 cluster

Cluster	DBI	Cluster yang dihasilkan	Anggota Cluster
2	0,667	2	Cluster 1 : 62 &

	1	T	1
			Cluster 2:
	0.702	2	
3	0,782	3	Cluster 1:
			57, Cluster
			2 : 7 &
			Cluster 3:
			62
4	1,150	4	Cluster 1 :
7	1,130	7	
			49, Cluster
			2: 21,
			Cluster 3:
			10& Cluster
			4:36
5	1,629	5	Cluster 1:
	,		49, Cluster
			2 : 21,
			Cluster 3 :
			10, Cluster
1 .			4 : 36&
JH_{A}			Cluster 5:7
6	0,625	6	Cluster 1:
			12, Cluster
			2 : 21,
			Cluster 3:
			35, Cluster
3 11/1/			
/			1 . 7
WE .			4 : 7,
			Cluster 5:
	* 3		Cluster 5: 19& Cluster
	X X		Cluster 5: 19& Cluster 6: 29
7	0,947	7	Cluster 5: 19& Cluster
7	0,947	7	Cluster 5: 19& Cluster 6: 29
7	0,947	7	Cluster 5 : 19& Cluster 6 : 29 Cluster 1 :
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21,
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3:
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7,
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6:
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster
7		7	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21
7		8	Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10,
ER	*		Cluster 5 : 19& Cluster 6: 29 Cluster 1 : 10, Cluster 2 : 21, Cluster 3 : 4, Cluster 4 : 25, Cluster 5 : 7, Cluster 6 : 35& Cluster 7: 21 Cluster 1 : 12, Cluster 2 : 10, Cluster 3 :
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 3: 21, Cluster
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 4: 35,
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 4: 35, Cluster 4: 35, Cluster 5:
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 4: 35, Cluster 4: 35, Cluster 5: 7, Cluster 6
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 4: 35, Cluster 4: 35, Cluster 5: 7, Cluster 5: 7, Cluster 6: 9, Cluster 6: 9, Cluster 6: 9, Cluster
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 4: 25, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 4: 35, Cluster 4: 35, Cluster 5: 7, Cluster 6
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 4: 35, Cluster 4: 35, Cluster 5: 7, Cluster 5: 7, Cluster 6: 9, Cluster 7: 10&
ER	*		Cluster 5: 19& Cluster 6: 29 Cluster 1: 10, Cluster 2: 21, Cluster 3: 4, Cluster 5: 7, Cluster 6: 35& Cluster 7: 21 Cluster 1: 12, Cluster 2: 10, Cluster 3: 21, Cluster 4: 35, Cluster 4: 35, Cluster 5: 7, Cluster 5: 7, Cluster 6: 9, Cluster 6: 9, Cluster 6: 9, Cluster

9	0,908	9	Cluster 1:
			12, Cluster
			2 : 21,
			Cluster 3:
			13, Cluster
			4 : 10,
			Cluster 5:
			35, Cluster
			6 : 7,
			Cluster 7:
			9, Cluster 8
			: 4 &
			Cluster 9:
			12
10	1,174	10	Cluster 1:
10	1,174	10	Cluster 1: 6, Cluster 2
10	1,174	10	
10	1,174	10	6, Cluster 2: 19, Cluster 3: 14,
10	1,174	10	6, Cluster 2: 19, Cluster 3: 14, Cluster 4:
10	1,174	10	6, Cluster 2: 19, Cluster 3: 14, Cluster 4: 10, Cluster
10	1,174	10	6, Cluster 2 : 19, Cluster 3 : 14, Cluster 4: 10, Cluster 5 : 10,
10	1,174	10	6, Cluster 2 : 19, Cluster 3 : 14, Cluster 4: 10, Cluster 5 : 10, Cluster 6:
10	1,174	10	6, Cluster 2 : 19, Cluster 3 : 14, Cluster 4: 10, Cluster 5 : 10, Cluster 6: 1, Cluster 7
10	1,174	10	6, Cluster 2 : 19, Cluster 3 : 14, Cluster 4: 10, Cluster 5: 10, Cluster 6: 1, Cluster 7: 27, Cluster
10	1,174	10	6, Cluster 2 : 19, Cluster 3 : 14, Cluster 4: 10, Cluster 5 : 10, Cluster 6: 1, Cluster 7 : 27, Cluster 8 : 10,
10	1,174	10	6, Cluster 2 : 19, Cluster 3 : 14, Cluster 4 : 10, Cluster 5 : 10, Cluster 6 : 1, Cluster 7 : 27, Cluster 8 : 10, Cluster 9 :
10	1,174	10 X	6, Cluster 2 : 19, Cluster 3 : 14, Cluster 4: 10, Cluster 5 : 10, Cluster 6: 1, Cluster 7 : 27, Cluster 8 : 10,

2. Pada tabel dibawah ini adalah profilling perbandingan cluster fitur.

Tabel 4.3 Anggota cluster 1 data pasien DBD

	Cluster 1			O.E.
NO	Hemoglo	Leuko	Hemato	Tromb
	bin	sit	krit	osit
1.	12,6	7900	35	124000
24.	14,8	5100	42	127000
		•		
33.	14,2	2800	39	123000
57.	15,1	5600	42	128000
92.	14,2	2800	39	123000
108.	14,1	7500	39	139000

Hasil pada tabel 4.2 anggota cluster 1 terdapat 12 data pasien demam berdarah dengan parameter:

- 1. Hemoglobin tertinggi ada pada nilai 15,1 dan terendah 12,6.
- 2. Leukosit tertinggi ada pada nilai 7900 dan terendah 2800.
- 3. Hematokrit tertinggi ada pada nilai 42 dan terendah 35.
- 4. Trombosit tertinggi ada pada nilai 139000 dan terendah 123000.

Tabel 4.4 Anggota cluster 2 data pasien DBD

	Cluster 1			
NO	Hemoglo	Leuko	Hemato	Trombo
	bin	sit	krit	sit
34.	16,2	3800	44	15000
J., Z.,				
37.	14,3	2600	39,7	34000
39.	12,2	4600	36	43000
		<u>.</u>		
56.	12,6	10700	36	45000
72.	15,3	4500	43	49000
	**			
90.	17,8	6700	50	19000
5E.				
98.	12,2	4600	36	43000
118				
	16	5000	57	30000
••		•		

Hasil pada tabel 4.3 anggota cluster 2 terdapat 21 data pasien demam berdarah dengan parameter:

- 1. Hemoglobin tertinggi ada pada nilai 17,8 dan terendah 12,2.
- 2. Leukosit tertinggi ada pada nilai 10700 dan terendah 2600.

- 3. Hematokrit tertinggi ada pada nilai 57 dan terendah 36.
- 4. Trombosit tertinggi ada pada nilai 49000 dan terendah 15000.

Tabel 4.5 Anggota cluster 3 data pasien DBD

	Cluster 3			
NO	Hemoglo	Leuko	Hemato	Tromb
	bin	sit	krit	osit
12.	12,2	2500	36	144000
••				
20.	12,1	2200	35	155000
	•••••			
••		•		
21.	17	7400	48	182000
		•••••		
		•		
28.	15,9	9100	45	156000
			<i></i>	
••		. //		
30.	10,9	4500	32	180000
••			44	

Hasil pada tabel 4.4 anggota cluster 3 terdapat 35 data pasien demam berdarah dengan parameter:

- 1. Hemoglobin tertinggi ada pada nilai 17 dan terendah 10,9.
- 2. Leukosit tertinggi ada pada nilai 9100 dan terendah 2200.
- 3. Hematokrit tertinggi ada pada nilai 48 dan terendah 32.
- 4. Trombosit tertinggi ada pada nilai 182000 dan terendah 144000.

Tabel 4.6 Anggota cluster 4 data pasien DBD

	pasien E			
	Cluster 4			
NO	Hemoglo	Leuko	Hemato	Tromb
	bin	sit	krit	osit
2.	11,8	12800	35	306000
••••				
••				
18.	16,1	6300	44	242000
47.	14,4	6300	39	230000

•••		•		
61.	11,8	12800	35	306000
••		•		
107				
	13,9	6800	39	216000
		•		

Hasil pada tabel 4.5 anggota cluster 4 terdapat 7 data pasien demam berdarah dengan parameter:

- 1. Hemoglobin tertinggi ada pada nilai 16,1 dan terendah 11,8.
- 2. Leukosit tertinggi ada pada nilai 12800 dan terendah 6300.
- 3. Hematokrit tertinggi ada pada nilai 44 dan terendah 35.
- 4. Trombosit tertinggi ada pada nilai 306000 dan terendah 216000.

Tabel 4.7 Anggota cluster 5 data pasien DBD

11.	Cluster 5			
NO	Hemoglo	Leuko	Hemato	Tromb
	bin	sit	krit	osit
1	12,4	7400	34	86000
4.	14,9	4600	48	88000
5.	14,8	8300	42	88000
111		•		
	16,1	4100	47	70000
112				
•	14,5	2700	40	58000
117				
•	13,9	1400	39	77000

Hasil pada tabel 4.6 anggota cluster 5 terdapat 19 data pasien demam berdarah dengan parameter:

- 1. Hemoglobin tertinggi ada pada nilai 16,1 dan terendah 12,4
- 2. Leukosit tertinggi ada pada nilai 8300 dan terendah 1400.
- 3. Hematokrit tertinggi ada pada nilai 48 dan terendah 34.
- 4. Trombosit tertinggi ada pada nilai 88000 dan terendah 58000.

Tabel 4.8 Anggota cluster 6 data pasien DBD

	pasien L	עעע		
	Cluster 5			
NO	Hemoglo	Leuko	Hemato	Tromb
	bin	sit	krit	osit
3.	10,2	4500	32	116000
		•••••		
••				. 9
17.	13	9700	35	120000
••		///	5	
29.	15,9	11900	46	103000
			.1.1.1	
				3.3
42.	16,1	5900	45	105000
••			. 5	
54.	14,7	3800	41	96000
			\ %	

Hasil pada tabel 4.7 anggota cluster 6 terdapat 29 data pasien demam berdarah dengan parameter:

- 5. Hemoglobin tertinggi ada pada nilai 16.1 dan terendah 10.2.
- 6. Leukosit tertinggi ada pada nilai 11900 dan terendah 3800.
- 7. Hematokrit tertinggi ada pada nilai 46 dan terendah 32.
- 8. Trombosit tertinggi ada pada nilai 120000 dan terendah 96000.

3.analisis deskripsi setiap cluster

a. posisi hemoglobin di cluster 6 adalah terendah dibandingkan cluster lainnya, dan posisi hemoglobin di cluster 2 adalah tertinggi dibandingkan cluster lainnya.

- b. posisi leukosit di cluster 5 adalah terendah dibandingkan cluster lainnya, dan posisi leukosit di cluster 4 adalah tertinggi dibandingkan cluster lainnya.
- c. posisi hematokrit di cluster 3&6 adalah terendah dibandingkan cluster lainnya, dan posisi hematokrit di cluster 2 adalah tertinggi dibandingkan cluster lainnya.
- d. posisi trombosit di cluster 2 adalah terendah dibandingkan cluster lainnya, dan posisi trombosit di cluster 3 adalah tertinggi dibandingkan cluster lainnya.

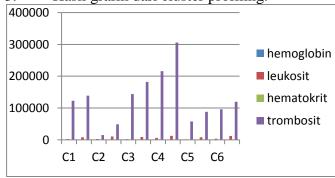
VIII. KESIMPULAN DAN SARAN

KESIMPULAN

Berdasarkan hasil penelitian dan pengujian yang telah dilakukan, dapat disimpulkan bahwa:

- 1. Didapatkan hasil cluster terbaik ada pada 6 cluster yang dihitung menggunakan *RapidMiner* dan didapatkan nilai DBI 0,625
- 2. Dari pengelompokan metode K-Medoids dihasilkan 6 cluster yaitu:
 - a. *Cluster* 1 terdapat 12 data pasien demam berdarah.
 - b. *Cluster* 2 terdapat 21 data pasien demam berdarah.
 - c. *Cluster* 3 terdapat 35 data pasien demam berdarah.
 - d. *Cluster* 4 terdapat 7 data pasien demam berdarah.
 - e. *Cluster* 5 terdapat 19 data pasien demam berdarah.
 - f. *Cluster* 6 terdapat 29 data pasien demam berdarah.

3. Hasil grafik dari cluster profiling.



SARAN

Beberapa saran yang dapat dijadikan dalam pengembangan penelitian ini adalah:

- 1. Pengembangan selanjutnyya disarankan dapatmembangun sebuah sistem/aplikasi yang dapat lebih membantu untuk proses pendiagnosaan pasien, karena penelitian ini hanya sebagai penerapan dari algoritma *K-Medoids*.
- 2. Untuk penentuan jumlah *cluster* terbaik bisa menggunakan metode yang lainnya tidak hanya metode DBI (*Davies Bouldin Index*)

DAFTARPUSTAKA

Evelyn C.P., 2009. Anatomi dan Fisiologi untuk paramedis. Jakarta. Gramedia.

Furqon M.T., Ridok. A., Mahmudy W.F., 2015. Parelisasi algoritma *K-Medoid* pada GPU menggunakan *open CL*.

Han, J., & Kamber, M. (2006). *Data Mining*: concepts and Techniques.

Kasasih E.N dan Kosasih A.S., 2008. Tafsiran hasil pemeriksaan labolatorium klinik edisi kedua.

Kementrian Kesehatan RI, 2010. Rencana Strategis kementrian kesehatan tahun 2010-2014. Jakarta.

Larose D,T., 2005. Discovering knowledge in data: an introduction to data mining, jhon Wiley & Sons Inc.

Pramudio, I. 2006. Apa itu data mining? Dalam

http://datamining.japati.net/cgibin/indodm.cgi
. Diakses tanggal 26 februari 2011.

Soegijanto S., 2004. Epidemiologi *demam* berdarah dengue. Dalam: Soegeng Soegijanto: Demam berdarah dengue: tinjauan dan temuan baru di era 2003. Surabaya Airlangga University press.

Soegijanto S., 2006 . Demam Berdarah Dengue. Edisi kedua. Surabaya : Airlangga University press.

Satari H.I. dan Meiliasari M. 2004. *Demam berdarah* . jakarta : Puspa Swara.

Sutedjo, A.Y. 2006. Mengenal penyakit melalui pemeriksaan labolatorium, Yogyakarta: Amara Books.

Santoso, S.2010. *statistik Multivariat*. Jakarta: Elex Media Komputindo.

Setyawati, A.W., 2017. Implementasi Algoritma *partitioning ArroundMedoid* (PAM) untuk pengelompokkan sekolah menengah atas di DIY berdasarkan nilai daya serap Ujian Nasional. Skripsi. Yogyakarta: Fakultas Sains dan Teknologi Universitas Sanata Dharma.

Tjay T.H. dan Rahardja K., 2006. "TAHAPAN KNOWLEDGE DISCOVERY in DATABASE". Tjay, Tan Hoan dan Kirana Rahardja ., 2007. Obat-obat penting khasiat, penggunaan dan efek-efek sampingnya.

Turban, E. dan Aronson. J.E., 2005. Decision Support System and *Intelligent System*. Yogyakarta: Andi Offset.

https://www.statistikian.com/2014/03/analisis-cluster.html