

Deteksi Penyebaran Penyakit Tuberkulosis dengan Algoritma *K-Means Clustering* Menggunakan *Rapid Miner*

Prasetyo Arta Kusuma^{1*)}, Ada Udi Firmansyah²⁾

¹⁾²⁾ Teknik Informatika, STMIK Dharma Wacana

^{*)}Correspondence Author: prasetyoarta02@gmail.com, Indonesia

DOI: <https://doi.org/10.37012/jtik.v8i2.1173>

Abstrak

Tuberkulosis (TBC) ialah penyakit yang diakibatkan oleh infeksi organisme Mikroskopis *Mycobacterium tuberculosis* yang menjangkiti bagian organ paru-paru. Ditahun 2016 ada 10,4 juta kasus TBC di planet ini, identik dengan 120 kasus untuk setiap 100.000 jiwa. Cina, India, Filipina, Pakistan, dan Indonesia adalah negara dengan kasus paling penting. Sebagian besar kejadian TBC yang dinilai pada tahun 2016 terjadi di Lokal Asia Tenggara (45%) dimana Indonesia nomor satu dan 25% terjadi di kawasan Afrika. Indonesia memiliki masalah yang besar dalam menangani penyakit TBC. Pada tahun 2017 terdapat 420.994 kasus TBC di Indonesia. Di Indonesia, penyakit TBC hampir menyebar di semua wilayah, salah satunya daerah Kota Bandung, Wilayah Jawa Barat. Ditahun 2020 terdapat 10758 kasus, dan yang paling terdampak adalah Wilayah Bojongloa Kaler sebanyak 879 kasus. Mengingat jumlah kasus TBC di Kota Bandung yang terus meningkat, dapat dibayangkan bahwa diperlukan upaya pengobatan yang kuat. Penelitian ini bertujuan untuk menentukan wilayah penyebaran penyakit TBC di Kota Bandung pada tahun 2020, penelitian kuantitatif ini menggunakan data yang diperoleh dari *Website Portal Data Kota Bandung* sebesar 10758 kasus penyakit TBC. Metode *Algoritma* yang digunakan ialah *K-Means* dan diolah menggunakan *Software Rapidminer*. Berdasarkan hasil uji dapat ditarik kesimpulan bahwa penyebaran anggota *cluster* terbanyak ada di *cluster* 0 dengan 10 anggota, lalu penyebaran yang berukuran sedang ada di *cluster* 4 dengan 8 anggota, untuk yang terkecil berada di *cluster* 3 dengan 1 anggota.

Kata Kunci : Tuberkulosis, *Data mining*, *Rapid Miner*, *Clustering*, *K-Means*

Abstract

Tuberculosis (TBC) is a disease caused by infection with the microscopic organism Mycobacterium tuberculosis that infects parts of the lungs. In 2016 there were 10.4 million TBC cases on the planet, identical to 120 cases for every 100,000 people. China, India, the Philippines, Pakistan and Indonesia are the countries with the most important cases. Most of the TBC incidence assessed in 2016 occurred in Local Southeast Asia (45%) where Indonesia was number one and 25% occurred in the African region. Indonesia has a big problem in dealing with TBC disease. In 2017 there were 420,994 TBC cases in Indonesia. In Indonesia, TBC disease has almost spread in all regions, one of which is Bandung City, West Java Region. In 2020 there were 10758 cases, and the most affected was the Bojongloa Kaler Region with 879 cases. Given the increasing number of TBC cases in the city of Bandung, it is conceivable that strong treatment efforts are needed. This study aims to determine the area of the spread of TBC disease in the city of Bandung in 2020, this quantitative study uses data obtained from the Bandung City Data Portal Website of 10758 cases of TBC disease. The algorithm method used is K-Means and processed using Rapidminer Software. Based on the test results, it can be concluded that the largest distribution of cluster members is in cluster 0 with 10 members, then the medium-sized distribution is in cluster 4 with 8 members, for the smallest one is in cluster 3 with 1 member.

Keywords: Tuberculosis, *Data mining*, *Rapid Miner*, *Clustering*, *K – Means*

PENDAHULUAN

Tuberkulosis (TBC) ialah penyakit yang diakibatkan oleh infeksi organisme Mikroskopis *Mycobacterium tuberculosis* yang menjangkiti bagian organ paru-paru (Torsa, 2020). Ditahun 2016 ada 10,4 juta kasus TBC di planet ini, identik dengan 120 kasus untuk

setiap 100.000 jiwa. Cina, India, Filipina, Pakistan, dan Indonesia adalah negara dengan kasus paling penting. Sebagian besar kejadian TBC yang dinilai pada tahun 2016 terjadi di Lokal Asia Tenggara (45%) dimana Indonesia nomor satu dan 25% terjadi di kawasan Afrika. Indonesia memiliki masalah yang besar dalam menangani penyakit TBC, pada tahun 2017 terdapat 420.994 kasus TBC di Indonesia (informasi per 17 Mei 2018) (Kemenkes RI-2018). Pada Indonesia, penyakit TBC hampir menyebar di semua wilayah, salah satunya daerah Kota Bandung, Wilayah Jawa Barat. Ditahun 2020 terdapat 10758 kasus, dan kasus yang paling terdampak adalah Wilayah Bojongloa Kaler sebanyak 879 kasus (Dinas Kesehatan Kota Bandung 2020). Mengingat jumlah kasus TBC di Kota Bandung yang terus meningkat, dapat dibayangkan bahwa diperlukan upaya pengobatan yang kuat. Secara khusus mengenali penyebaran penyakit TBC menurut sub wilayah di Kota Bandung. Dengan demikian, identifikasi penyebaran penyakit TBC menurut sub-wilayah diperlukan untuk menentukan kualitas *cluster* sub-lokal mana yang paling mendekati kasus TBC tinggi dan rendah. Karena penentuan *cluster* ini bisa menjadi salah satu kunci penanggulangan penyakit TBC.

Untuk menentukan kualitas *cluster* disuatu wilayah, diperlukannya sebuah metode yang akurat. Beberapa metode algoritma yang bisa digunakan yaitu *K-Means*, *Mixture Modelling*, *Self-Organising Map* (SOM) dan masih banyak lagi. Metode *K-Means* yakni satu diantara banyak metode yang dipakai dalam kegiatan klusterisasi. Metode ini efektif menghasilkan *cluster-cluster* didalam tumpukan data.

(Toresa, 2020) dalam penelitiannya tentang “Implementasi *K-Means* Terhadap Penyebaran Penyakit TBC di Riau Menggunakan *Rapid Miner*”. Dimana menurutnya pemilihan *cluster* terbaik ialah menggunakan metode *K-Means Clustering* dan tidak menggunakan pengukuran *cluster optimum*. Data kasus Tuberkulosis di Provinsi Riau didapat dari diskusi langsung dengan Badan Pusat Statistik Provinsi Riau, lalu data tersebut diolah menggunakan *Software Rapid Miner*. Hasil penelitian adalah ditemukannya 3 daerah dengan level *cluster* rendah, ada 7 daerah dengan level *cluster* sedang, dan ada 2 daerah memiliki level *cluster* tinggi, dari 12 wilayah Provinsi Riau.

Dalam penelitian (Widianingrum et. Al, 2020) ditemukan bahwa implementasi data mining memakai Metode Algoritma *K-Means Clustering* terbukti efektif untuk membuktikan pola transmisi TBC di Kabupaten Ponorogo dimana *Software Rapidminer* digunakan untuk memproses datanya dan berhasil ditemukan daerah mana saja yang memiliki *history* transmisi TBC. Dari hasil penelitian dapat disimpulkan bahwa transmisi TBC terbagi dalam beberapa kategori seperti tinggi, dipertimbangkan, dan siaga. Adapun

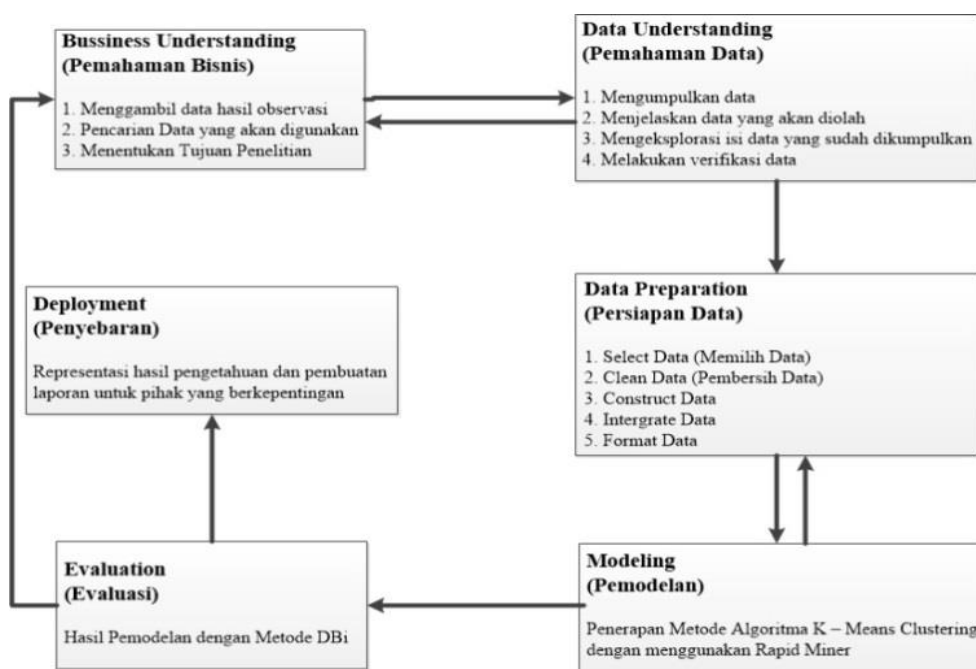
beberapa daerah yang dikategorikan dengan tingkat transmisi tinggi berada di Puskesmas Ngebel, dan Puskesmas Babadan. Untuk tingkat transmisi TBC dengan kategori transmisi dipertimbangkan dapat ditemukan di Puskesmas Nailan. Untuk pengembangan penelitian lebih lanjut dapat menggunakan data penderita penyakit TBC yang tersebar di Kabupaten Ponorogo. Untuk memantau penanganan penderita TBC dibutuhkan pendataan terintegrasi dan penanganan yang serius.

Dari pokok permasalahan tersebut maka dilakukan penelitian yang berjudul “Deteksi Penyebaran Penyakit Tuberkulosis dengan Algoritma *K-Means Clustering* menggunakan *Rapid Miner*” untuk mengelompokkan Kecamatan di Kota Bandung berdasarkan data TBC tahun 2020. Penelitian yang dilaksanakan menggunakan Metode Algoritma *K-Means Clustering* memakai metode DBi (*Davies Bouldin Index*) digunakan untuk penentuan *cluster* optimum dalam menentukan *cluster* terbaik. Lalu dilakukan uji coba sebanyak 2 hingga 10 *Clustering* untuk menentukan *cluster* optimumnya. Hasil dari penelitian ini nantinya bisa jadi bahan usulan bagi Pemerintah Daerah Bandung dalam menyelesaikan pengobatan penyakit TB yang lebih ampuh. Pemerintah diharapkan bisa lebih fokus terlebih dulu ke wilayah-wilayah yang memiliki permasalahan TBC tingkat besar, sehingga jumlah permasalahan TBC di Kecamatan Kota Bandung terus menurun.

METODE

Cross Industry Standard Process Model for Data mining (CRISP-DM) mempunyai 6 proses, yaitu (*Business Understanding*) Pemahaman Bisnis, (*Data Understanding*) Pemahaman Data, (*Data Preparation*) Persiapan Data, (*Modeling*) Pemodelan, (*Evaluation*) Evaluasi, dan Penyebaran (*Deployment*). Metode Algoritma *K-Means Clustering* dipakai dalam penelitian ini untuk mengelompokkan data persoalan TBC di Kota Bandung tahun 2020 tanpa membuat suatu program aplikasi. Untuk data dan informasi bersumber dari Website Portal Data Kota Bandung. Untuk pengolahan data dilakukan dengan menggunakan *Software Rapid Miner* lalu dilakukan penentuan jumlah *cluster* yang nantinya dibuat dengan memakai Metode *Davies Bouldin index* (DBi).

Tahap *Business Understanding* (Pemahaman Bisnis) ialah cara untuk memastikan alasan penelitian. Penelitian ini bertujuan untuk pengelompokan wilayah penyebaran TBC di Kota Bandung dengan memakai Metode Algoritma *K-Means Clustering*. Hasil Observasi yang diambil dari *Website* Portal Data Kota Bandung yaitu mempersiapkan data yang akan digunakan untuk diolah. Dari pengolahan data tersebut nantinya dapat membantu pemerintah Kota Bandung dalam menyelesaikan penanganan kasus penyakit TB.



Gambar 1. Tahap-tahap CRIPS – DM

Data Understanding atau Pemahaman Data tentang masalah TBC di Kecamatan Kota Bandung tahun 2020. Dibawah ini adalah gambar tabel deskripsi data tersebut :

Tabel 1. Deskripsi Data Penyakit TBC

Attribut	Tipe Data	Keterangan
Kecamatan	Varchar	Kecamatan di Kota Bandung
Jumlah Kasus Baru TB BTA+	Integer	Pengidap baru telah cek lab BTA+
Jumlah Kasus TB Anak	Integer	Kasus TB Anak di Wilayah Kecamatan

Tabel 1 adalah penjelasan mengenai data TB pada Kota Bandung tahun 2020. Data tersebut menjelaskan attribut, tipe data dan karakteristik data.

Tabel 2. Penjelasan Data Penyakit TBC

Attribut	Tipe Data	Keterangan
Kecamatan	Varchar	Kecamatan di Kota Bandung Tahun 2020 : 30 Kecamatan
Jumlah Kasus Baru TB BTA+	Integer	Pengidap baru telah cek lab BTA+ Tahun 2020 : 6757 Pasien
Jumlah Kasus TBC Anak	Integer	Kasus TBC Anak di Wilayah Kecamatan Tahun 2020 : 4001 Pasien

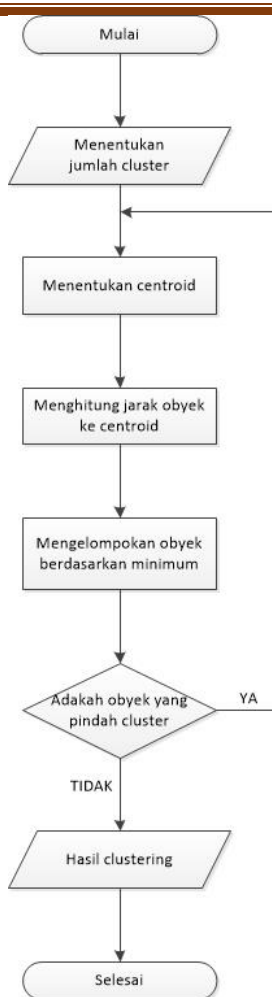
Tabel 2 adalah penjelasan secara rinci tentang banyaknya jumlah dari atribut.

Selanjutnya mempersiapkan (*Data Preparation*) data kasus TBC agar bisa menjadi data final / dataset. Data tersebut tidak perlu dibersihkan, karena saat proses pemodelan data, data tersebut tidak *missing* dan tidak memerlukan normalisasi, karena tipe data yang sama.

Tabel 3. Data Final Penyakit TBC

Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak
Andir	196	140
Antapani	225	135
Arcamanik	92	76
Astanaanyar	336	260
Babakan Ciparay	495	309
Bandung Kidul	210	87
Bandung Kulon	279	225
Bandung Wetan	72	26
Batununggal	360	198
Bojongloa Kaler	498	381
Bojongloa Kidul	333	144
Buahbatu	214	84
Cibeunying Kaler	120	78
Cibeunying Kidul	224	114
Cibiru	216	105
Cicendo	134	90
Cidadap	90	38
Cinambo	21	16
Coblong	488	212
Gedebage	72	52
Kiaracondong	298	152
Lengkong	288	156
Mandalajati	290	160
Panyileukan	68	42
Rancasari	108	78
Regol	180	117
Sukajadi	300	186
Sukasari	272	196
Sumur Bandung	94	54
Ujungberung	184	90

Berikut ini adalah cara perhitungan Metode Algoritma *K-Means*.



Gambar 2. Flowchart Alur Metode *Algoritma K – Means*

Data Penyakit TB di Kota Bandung tahun 2020 terdiri dari 30 Kecamatan, lalu di proses menggunakan Metode *Algoritma K-Means* dalam pengelompokan kecamatan berdasarkan kasus.

Tabel 4. Perhitungan Manual Metode *Algoritma K – Means*

Kecamatan	Centroid 1	Centroid 2	Centroid 3	Centroid 4	Cluster			
Babakan Ciparay	8,439326	225,1314	480,5189	351,3528	1			
Bojongloa Kaler	80,45012	273,2651	524,5139	398,9674	1			
Coblong	88,84756	183,5809	430,5459	300,1042	1			
Astanaanyar	162,8268	79,4575	322,4534	199,1696		2		
Bandung Kulon	227,612	47,28806	256,3701	136,9918		2		
Batununggal	168,5444	55,02872	308,0964	177,7597		2		
Bojongloa Kidul	224,4064	50,09152	261,5106	131,6141		2		
Kiaracondong	245,7381	35,30415	232,1375	101,4397		2		
Lengkong	251,4502	35,38588	224,8595	94,40612		2		
Mandalajati	247,5222	30,92903	228,4588	98,16321		2		
Sukajadi	225,0672	6,231144	249,9748	121,4146		2		
Sukasari	245,1351	35,56128	232,5274	109,1261		2		
Arcamanik	460,2295	240,966	21,56409	118,8003				3
Bandung Wetan	503,2341	283,843	32,69572	157,7292				3

Cibeunying Kaler	434,9796	215,441	40,14237	91,53423			3			
Cicendo	416,822	197,3337	58,52017	74,58563			3			
Cidadap	481,6021	262,2114	17,24558	136,1103			3			
Cinambo	551,7689	332,2125	76,7477	207,1721			3			
Gedebage	489,5282	270,0102	15,39513	145,7344			3			
Panyileukan	498,0969	278,5353	23,10433	153,5172			3			
Rancasari	445,3305	225,8941	31,07748	102,9054			3			
Sumur Bandung	469,6572	250,1007	6,972087	124,888			3			
Andir	338,2591	119,5647	138,1456	32,61159				4		
Antapani	315,6378	96,08413	159,4252	32,12889				4		
Bandung Kidul	355,1341	138,2962	126,9977	22,33866				4		
Buahbatu	353,7766	137,7572	130,1715	26,21098				4		
Cibeunying Kidul	327,9704	109,5108	149,0725	18,56113				4		
Cibiru	339,6825	121,4708	138,2578	10,65437				4		
Regol	363,4835	144,011	111,6889	27,32244				4		
Ujungberung	374,5316	155,6226	103,0272	29,1636				4		
Total Data Cluster							3	9	10	8

Tabel 4 adalah hasil perhitungan manual Metode Algoritma *K-Means* menggunakan *Software Microsoft Excel*. Data tersebut dihitung ulang sampai dengan iterasi ke 8, karena di iterasi ke 8 posisi *cluster* sudah tidak berpindah ke *cluster* lainnya.

Tahap evaluasi adalah perhitungan dengan metode validasi yaitu *Davies Bouldin Index* (DBi) untuk mendapatkan nilai *cluster* yang terbaik. *Cluster* yang akan di hitung yaitu $K=4$, nilai Dbi dari $K=4$ adalah 0,50417.

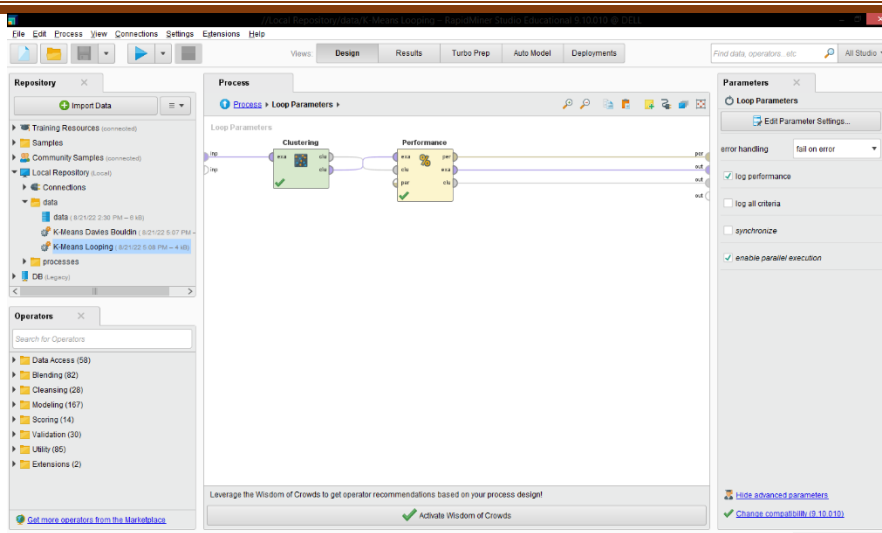
Tabel 5. Hasil perhitungan DBi

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Max	DBi
1	0	0,459748	0,192815	0,526295	0,526295	0,50417
2	0,459748	0	0,289838	0,526295	0,526295	
3	0,192815	0,289838	0	0,437794	0,437794	
4	0,243425	0,526295	0,437794	0	0,526295	

Tahap yang terakhir adalah penyebaran informasi data yang sudah dibuat dalam bentuk *data mining*. Tujuannya agar Dinas Kesehatan Kota Bandung dapat menangani kasus dan permasalahan penyakit TBC di Kota Bandung.

HASIL DAN PEMBAHASAN

Data yang dipakai diambil dari Website Portal Data Kota Bandung, yaitu Data Penyakit TBC di Kota Bandung tahun 2020 dengan 2 attribut Kasus Baru BTA+, dan Kasus TBC Anak. Lalu melakukan pemrosesan penerapan *cluster* menggunakan Metode *Algoritma K-Means* dengan memakai *Software Rapid Miner*, berikut adalah gambarannya :



Gambar 3. Proses *Clustering* metode *K-Means*

Hasil perhitungan *cluster* dengan nilai DBi terkecil menggunakan *Rapid Miner* dapat dilihat pada tabel 6 berikut ini.

Tabel 6. Hasil perhitungan *cluster* dengan nilai DBi dengan anggota *cluster* nya

<i>Clustering. K</i>	Nilai DBi	Anggota Cluster
2	0,641	<i>Cluster_0</i> : 18 items, <i>Cluster_1</i> : 12 items
3	0,619	<i>Cluster_0</i> : 5 items, <i>Cluster_1</i> : 10 items, <i>Cluster_2</i> : 15 items
4	0,488	<i>Cluster_0</i> : 9 items, <i>Cluster_1</i> : 10 items, <i>Cluster_2</i> : 3 items, <i>Cluster_3</i> : 8 items
5	0,415	<i>Cluster_0</i> : 10 items, <i>Cluster_1</i> : 2 items, <i>Cluster_2</i> : 9 items, <i>Cluster_3</i> : 1 items, <i>Cluster_4</i> : 8 items
6	0,531	<i>Cluster_0</i> : 2 items, <i>Cluster_1</i> : 10 items, <i>Cluster_2</i> : 7 items, <i>Cluster_3</i> : 8 items, <i>Cluster_4</i> : 1 items, <i>Cluster_5</i> : 2 items
7	0,600	<i>Cluster_0</i> : 1 items, <i>Cluster_1</i> : 4 items, <i>Cluster_2</i> : 6 items, <i>Cluster_3</i> : 2 items, <i>Cluster_4</i> : 3 items, <i>Cluster_5</i> : 8 items, <i>Cluster_6</i> : 6 items
8	0,571	<i>Cluster_0</i> : 5 items, <i>Cluster_1</i> : 1 items, <i>Cluster_2</i> : 2 items, <i>Cluster_3</i> : 2 items, <i>Cluster_4</i> : 6 items, <i>Cluster_5</i> : 8 items, <i>Cluster_6</i> : 5 items, <i>Cluster_7</i> : 1 items
9	0,511	<i>Cluster_0</i> : 4 items, <i>Cluster_1</i> : 2 items, <i>Cluster_2</i> : 3 items, <i>Cluster_3</i> : 5 items, <i>Cluster_4</i> : 8 items, <i>Cluster_5</i> : 2 items, <i>Cluster_6</i> : 1 items, <i>Cluster_7</i> : 4 items, <i>Cluster_8</i> : 1 items
10	0,455	<i>Cluster_0</i> : 1 items, <i>Cluster_1</i> : 1 items, <i>Cluster_2</i> : 4 items, <i>Cluster_3</i> : 1 items, <i>Cluster_4</i> : 5 items, <i>Cluster_5</i> : 8 items, <i>Cluster_6</i> : 2 items, <i>Cluster_7</i> : 3 items, <i>Cluster_8</i> : 1 items, <i>Cluster_9</i> : 4 items

Setelah dilakukannya pengujian *Clustering* dengan K = 2 sampai K = 10, nilai DBi yang terkecil adalah K = 5, dengan nilai 0,415. Setelah mendapat jumlah *cluster* yang optimal, selanjutnya menjelaskan anggota dan karakteristik yang terbentuk di cluster K = 5.

Tabel 7. Anggota *Cluster_0*

Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak	Cluster
Arcamanik	92	76	Cluster_0
Bandung Wetan	72	26	
Cibeunying Kaler	120	78	
Cicendo	134	90	
Cidadap	90	38	
Cinambo	21	16	
Gedebage	72	52	
Panyileukan	68	42	
Rancasari	108	78	
Sumur Bandung	94	54	

Tabel 7 adalah daftar nama-nama anggota *Cluster_0* dari hasil perhitungan menggunakan *Rapid Miner*.

Tabel 8. Karakteristik *Cluster_0*

<i>Cluster</i> : 0				
Jumlah Anggota : 10				
Karakteristik	Cluster 0	Kecamatan	Jumlah Kasus Baru TBC BTA +	Jumlah Kasus TBC Anak
	Kasus Tertinggi	Cicendo	134	90
	Kasus Terendah	Cinambo	21	16

Tabel 8 adalah nama anggota *Cluster_0* yang mempunyai karakteristik kasus tertinggi dan kasus terendah dari 10 anggota.

Tabel 9. Anggota *Cluster_1*

Kecamatan	Jumlah Kasus Baru TBC BTA +	Jumlah Kasus TBC Anak	Cluster
Babakan Ciparay	495	309	Cluster_1
Bojongloa Kaler	498	381	

Tabel 9 adalah daftar nama-nama anggota *Cluster_1* dari hasil perhitungan menggunakan *Rapid Miner*.

Tabel 10. Karakteristik *Cluster_1*

<i>Cluster</i> : 1				
Jumlah Anggota : 2				
Karakteristik	Cluster 1	Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak
	Kasus Tertinggi	Bojongloa Kaler	498	381
	Kasus Terendah	Babakan Ciparay	495	309

Tabel 10 adalah nama anggota *Cluster_1* yang mempunyai karakteristik kasus tertinggi dan kasus terendah dari 2 anggota.

Tabel 11. Anggota *Cluster_2*

Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak	Cluster
Astanaanyar	336	260	Cluster_2
Bandung Kulon	279	225	
Batununggal	360	198	
Bojongloa Kidul	333	144	
Kiaracondong	298	152	
Lengkong	288	156	
Mandalajati	290	160	
Sukajadi	300	186	
Sukasari	272	196	

Tabel 11 adalah daftar nama nama anggota *Cluster_2* dari hasil perhitungan menggunakan *Rapid Miner*.

Tabel 12. Karakteristik *Cluster_2*

<i>Cluster</i> : 2				
Jumlah Anggota : 9				
Karakteristik	Cluster 2	Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak
	Kasus Tertinggi	Batununggal	360	198
	Kasus Terendah	Sukasari	272	196

Tabel 12 adalah nama anggota *Cluster_2* yang mempunyai karakteristik kasus tertinggi dan kasus terendah dari 9 anggota.

Tabel 13. Anggota *Cluster_3*

Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak	Cluster
Coblong	488	212	Cluster_3

Tabel 13 adalah daftar nama nama anggota *Cluster_3* dari hasil perhitungan menggunakan *Rapid Miner*.

Tabel 14. Karakteristik *Cluster_3*

<i>Cluster</i> : 3				
Jumlah Anggota : 1				
Karakteristik	Cluster 3	Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak
	Kasus Tertinggi	Coblong	488	212
	Kasus Terendah	Coblong	488	212

Tabel 14 adalah nama anggota *Cluster_3* yang mempunyai karakteristik kasus tertinggi dan kasus terendah dari 1 anggota.

Tabel 15. Anggota *Cluster_4*

Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak	Cluster
Andir	196	140	Cluster_4
Antapani	225	135	
Bandung Kidul	210	87	
Buahbatu	214	84	
Cibeunying Kidul	224	114	
Cibiru	216	105	
Regol	180	117	
Ujungberung	184	90	

Tabel 15 adalah daftar nama nama anggota *Cluster_4* dari hasil perhitungan menggunakan *Rapid Miner*.

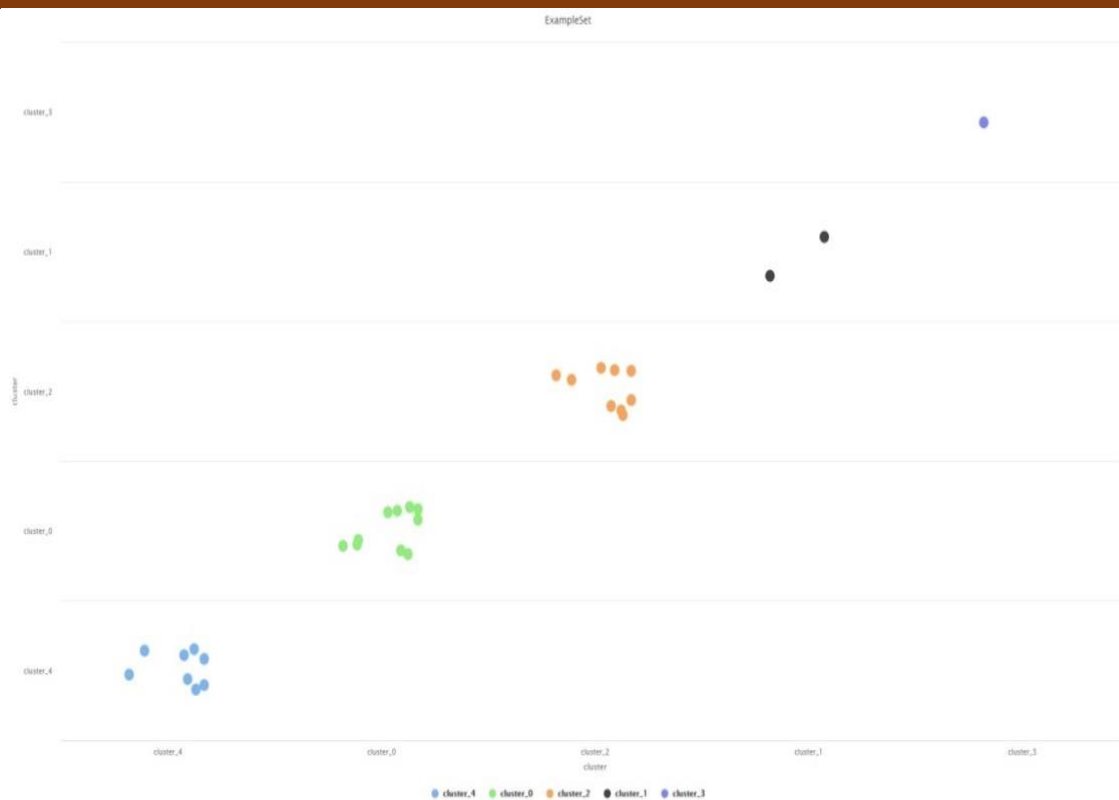
Tabel 16. Karakteristik *Cluster_4*

Cluster	: 4			
Jumlah Anggota	: 8			
Karakteristik	Cluster 2	Kecamatan	Jumlah Kasus Baru TB BTA +	Jumlah Kasus TB Anak
	Kasus Tertinggi	Antapani	225	135
	Kasus terendah	Regol	180	117

Tabel 16 adalah nama anggota *Cluster_4* yang mempunyai karakteristik kasus tertinggi dan kasus terendah dari 8 anggota.

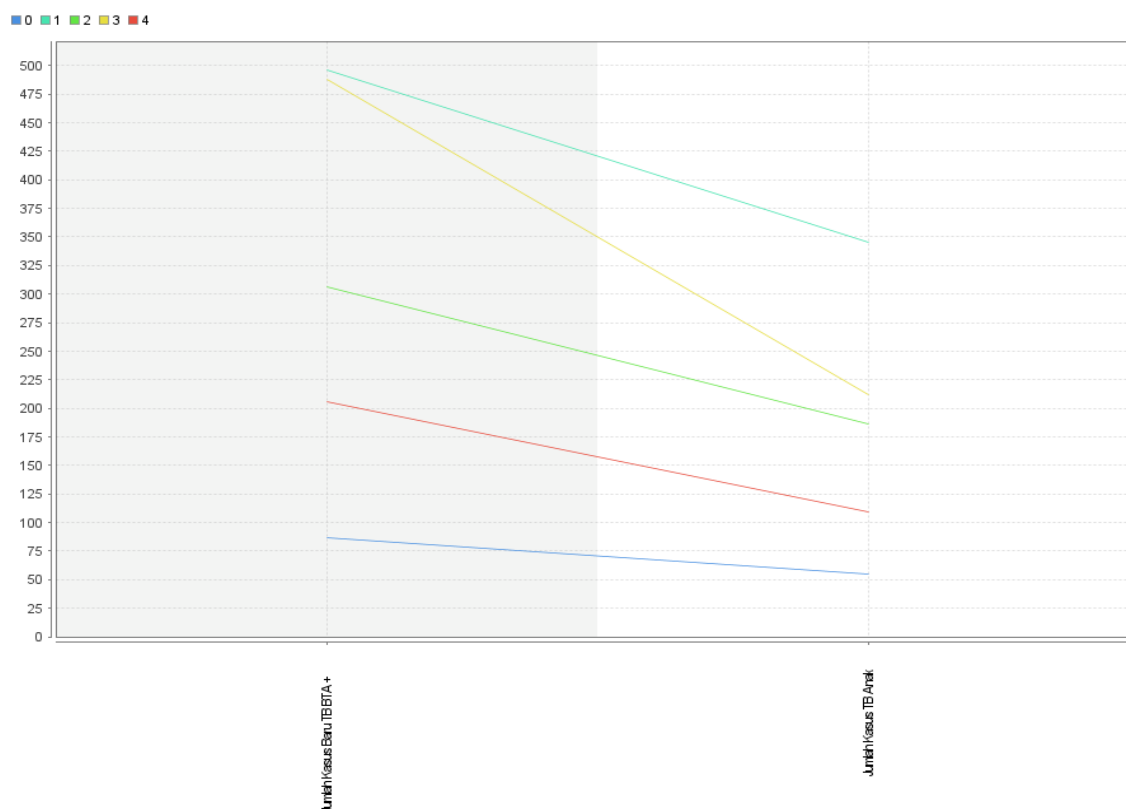
Telah dijelaskan anggota *cluster* dan karakteristik nya, selanjutnya dijelaskan karakteristik dari cluster yang terbentuk di setiap attribut yaitu Jumlah Kasus Baru TBC BTA+ dan Jumlah Kasus TBC anak pada cluster 2 mempunyai nilai paling tinggi dari *cluster* lainnya. Lalu untuk Jumlah Kasus Baru TBC BTA+ dan Jumlah Kasus TBC anak terendah berada pada *cluster* 0.

Gambar berikut menampilkan data penyebaran penyakit TBC di Kota Bandung dalam bentuk diagram *scatter*.



Gambar 4. Visualisasi Penyebaran Atribut Data

Terlihat pada Gambar 4 penyebaran anggota *cluster* terbanyak ada di cluster 0 dengan 10 anggota, lalu penyebaran yang berukuran sedang ada di *cluster* 4 dengan 8 anggota, untuk yang terkecil berada di *cluster* 3 dengan 1 anggota. Berikut ini adalah Gambar *Centroid Plot View*.



Gambar 5. Centroid Plot View

Pada Gambar 5 terlihat bahwa Jumlah Kasus Baru TBC BTA+ ternyata lebih tinggi dibandingkan dengan Jumlah Kasus TBC Anak.

KESIMPULAN DAN REKOMENDASI

Kesimpulan atau hasil dari penelitian Penyebaran Penyakit TBC menggunakan metode Algoritma *K-Means Clustering* dengan Nilai DBi ialah :

1. Penerapan metode *Algoritma K – Means Clustering* dalam penyebaran penyakit TBC di Kota Bandung mendapatkan hasil *cluster* optimum yaitu 5 *cluster* berdasarkan nilai DBi 0,415 dengan pengujian 2 sampai 10 *cluster*.
2. Anggota dari *cluster* optimum yaitu *Cluster 0* ada 10 Kecamatan, *Cluster 1* ada 2 Kecamatan, *Cluster 2* ada 9 Kecamatan, *Cluster 3* ada 1 Kecamatan, *Cluster 4* ada 8 Kecamatan.
3. Karakteristik dari *cluster* yang terbentuk di setiap atribut yaitu Jumlah Kasus Baru TBC BTA+ dan Jumlah Kasus TBC anak pada *cluster 2* mempunyai nilai paling tinggi dari *cluster* lainnya. Lalu untuk Jumlah Kasus Baru TBC BTA+ dan Jumlah Kasus TBC anak terendah berada pada *cluster 0*.

4. Penyebaran anggota *cluster* terbanyak ada di *cluster* 0 dengan 10 anggota, lalu penyebaran yang berukuran sedang ada di *cluster* 4 dengan 8 anggota, untuk yang terkecil berada di *cluster* 3 dengan 1 anggota.

REFERENSI

- Al-Rizki, Muhammad Farid Iqbal, Ida Widaningrum, and Ghulam Asrofi Buntoro. 2020. "Prediksi Penyebaran Penyakit TBC Dengan Metode *K-Means Clustering* Menggunakan Aplikasi Rapidminer." *JTERA (Jurnal Teknologi Rekayasa)* 5(1): 1.
- Algoritma, Penerapan et al. 2022. "Jurnal Smart Teknologi Jurnal Smart Teknologi." 3(3): 273–85.
- Fatmawati, Kiki, and Agus Perdana Windarto. 2018. "Data mining: Penerapan Rapidminer Dengan *K-Means Cluster* Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi." *Computer Engineering, Science and System Journal* 3(2): 173.
- Hartati, Tuti, Odi Nurdiawan, and Eko Wiyandi. 2021. "Analisis Dan Penerapan Algoritma *K-Means* Dalam Strategi Promosi Kampus Akademi Maritim Suaka Bahari." *Jurnal Sains Teknologi Transportasi Maritim* 3(1): 1–7.
- Hermanto, Teguh Iman et al. 2021. "Analisis Sebaran Titik Rawan Bencana Dengan *K-Means Clustering* Dalam Penanganan Bencana." *Jurnal Sains Komputer & Informatika (J-SAKTI)* 5(1): 406.
- Toresa, Dafwen. 2020. "Implementasi *K-Means* Terhadap Penyebaran Penyakit Tbc Di Riau Menggunakan *Rapid Miner*." *JUTIM (Jurnal Teknik Informatika Musirawas)* 5(1): 35–42.