

# Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Penyakit Tbc (Studi Kasus: Puskesmas Di Kabupaten Jember)

*by Nilogiri Agung*

---

**Submission date:** 10-Jun-2022 10:05AM (UTC+0800)

**Submission ID:** 1853967240

**File name:** baran\_Penyakit\_Tbc\_Studi\_Kasus\_Puskesmas\_Di\_Kabupaten\_Jember.pdf (1.93M)

**Word count:** 4446

**Character count:** 27372

**Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Penyakit Tbc  
(Studi Kasus: Puskesmas Di Kabupaten Jember)**

*Application Of K-Means Clustering Algorithm For Clustering The Spread Of Tb ( Case Study: Puskesmas  
In Jember Regency)*

**Bobby Purwanto<sup>1</sup>, Agung Nilogiri<sup>2\*</sup>, Ari Eko Wardoyo<sup>3</sup>**

<sup>1</sup>Mahasiswa Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

Email: [bobbypurwanto145@gmail.com](mailto:bobbypurwanto145@gmail.com)

<sup>2</sup>Dosen Fakultas Teknik, Universitas Muhammadiyah Jember \*Koresponden Author

Email: [agungnilogiri@unmuhjember.ac.id](mailto:agungnilogiri@unmuhjember.ac.id)

<sup>3</sup>Dosen Fakultas Teknik, Universitas Muhammadiyah Jember

Email: [ariek@unmuhjember.ac.id](mailto:ariek@unmuhjember.ac.id)

**Abstrak**

TBC (*Tuberculosis*) merupakan penyakit yang menyerang paru-paru yang ditimbulkan oleh kuman *Mycobacterium tuberculosis*. Di Indonesia penyakit TBC hampir tersebar di seluruh wilayah yang salah satunya di Kabupaten Jember Provinsi Jawa Timur dimana jumlah kasus TBC terus meningkat setiap tahunnya. Diperlukan sebuah upaya untuk mengelompokkan penyebaran penyakit TBC berdasarkan puskesmas di kabupaten Jember. Penelitian ini bertujuan untuk mencari cluster yang memiliki nilai optimum dan untuk mengetahui kelompok puskesmas yang terdapat dalam masing-masing cluster. Metode yang dapat digunakan untuk mengelompokkan daerah penyebaran penyakit TBC adalah data mining. Tahapan penelitiannya yang diterapkan yaitu CRISP-DM. Dataset yang digunakan bersumber dari Dinas Kesehatan Kabupaten Jember tahun 2017 - 2019. Kemudian data tersebut diproses menggunakan metode K-Means dan Davies Bouldin Index untuk menentukan cluster optimum dengan bantuan tools Rapid miner. Dari hasil penelitian ini diperoleh cluster optimum pada 9 cluster dengan nilai Davies Bouldin Index sebesar 0,784. Berdasarkan karakteristik cluster optimum menunjukkan bahwa kategori nilai kasus tingkat tinggi selama 3 tahun di semua atribut berada di cluster 4 berjumlah 1 anggota yaitu puskesmas Kalisat sedangkan pada Kategori nilai kasus tingkat rendah berada di cluster 2 berjumlah 12 anggota yaitu puskesmas Cakru, Tembokrejo, Lojejer, Sabrang, Curahnongko, Nogosari, Sukorejo, Klatakan, Jelbuk, Mangli, Ladakpakem, dan Banjarsengon.

**Kata Kunci:** TBC, *Data Mining, Clustering, CRISP-DM, K-Means.*

**Abstract**

*Tuberculosis* (TB) is a disease that attacks the lungs caused by the bacteria *Mycobacterium tuberculosis*. In Indonesia, TB disease is almost spread throughout the region, one of which is in Jember Regency, East Java Province, where the number of TB cases continues to increase every year. An effort is needed to classify the spread of TB disease based on puskesmas in Jember district. This study aims to find cluster that have the optimum value and to find out the groups of puskesmas contained in each cluster. The method that can be used to classify the area of spread of TB disease is data mining. The stages of the research that are applied are CRISP-DM. The dataset used is sourced from the Jember District Health Office in 2017 - 2019. Then the data is processed using the K-Means and Davies Bouldin Index methods to determine the optimum cluster with the help of Rapid miner tools. From the results of this study, the optimum cluster was obtained in 9 clusters with a Davies Bouldin Index value of 0.784. Based on the characteristics of the optimum cluster, it shows that the category of high-level case values for 3 years in all attributes is in cluster 4 with 1 member, namely Kalisat health center, while the low-level case value category is in cluster 2 with 12 members, namely

Cakru Health Center, Tembokrejo, Lojejer, Sabrang, Curahnongko, Nogosari, Sukorejo, Klatakan, Jelbuk, Mangli, **Madapakem**, and Banjarsengon.

**Keywords:** TB, Data Mining, Clustering, CRISP-DM, K-Means.

## 1. PENDAHULUAN

TBC (*Tuberculosis*) adalah penyakit yang menyerang paru-paru yang ditimbulkan oleh kuman *Mycobacterium tuberculosis* (Toresa, 2020). Pada tahun 2016 terdapat 10,4 juta kasus TBC di dunia, setara dengan 120 kasus per 100 ribu penduduk. India, Indonesia, China, Filipina, dan Pakistan merupakan negara dengan kasus tertinggi. Sebagian besar estimasi insiden TBC pada tahun 2016 terjadi di kawasan Asia Tenggara (45%) di mana Indonesia merupakan salah satu di dalamnya dan 25% terjadi di kawasan Afrika. Indonesia mempunyai masalah besar dalam penanganan penyakit TBC, di tahun 2017 kasus TBC di Indonesia terdapat 420.994 kasus (data per 17 Mei 2018) (KEMENKES RI-2018). Di Indonesia penyakit TBC nyaris tersebar di seluruh wilayah yang salah satunya ialah di Kabupaten Jember Provinsi Jawa Timur. Pada tahun 2017 sampai dengan 2019 orang yang terdampak penyakit TBC di Kabupaten Jember selalu meningkat. Terdapat 3853 kasus di tahun 2017, 3816 kasus pada tahun 2018, dan mengalami kenaikan yang signifikan di tahun 2019 yaitu terdapat 4527 kasus (DINKES Kabupaten Jember 2020). Karena jumlah kasus TBC di Kabupaten Jember terus meningkat dalam 3 tahun terakhir, memungkinkan perlunya suatu upaya penanganan yang efektif. Yaitu, dengan mengelompokkan penyebaran penyakit TBC berdasarkan puskesmas di kabupaten Jember. Maka dari itu, pengelompokan penyebaran penyakit TBC berdasarkan puskesmas bertujuan untuk mengetahui karakteristik puskesmas mana saja yang memiliki kemiripan terdekat dalam kasus TBC tingkat tinggi dan rendah. Hasil penelitian ini diharapkan dapat menjadi bahan rekomendasi untuk pemerintah kabupaten jember dalam melakukan penanganan yang lebih efektif terhadap penyakit TBC. Pemerintah nantinya dapat lebih fokus terlebih dahulu ke daerah – daerah yang memiliki kasus TBC tingkat tinggi, sehingga jumlah kasus TBC di puskesmas Kabupaten Jember dapat semakin berkurang.

Pada penelitian sebelumnya yang dilakukan oleh (Sari, Primajaya, & Irawan, 2020) dengan studi kasus “Implementasi Algoritma K-Means untuk Clustering Penyebaran Tuberculosis di Kabupaten Karawang” menggunakan data kasus Tuberculosis di Karawang tahun 2018. Penelitian yang dilakukan yaitu menggunakan software WEKA dan memanfaatkan metode elbow dalam pencarian jumlah cluster terbaik. Hasil uji coba yang dilakukan dengan menggunakan SSE (SUM of Square Error) dan Silhouette. Didapati K=3 sebagai cluster terbaik dan memanfaatkan kombinasi S=10. Hasilnya yaitu pada cluster 1 sebanyak 7 anggota, cluster 2 sebanyak 9 anggota, dan cluster 3 sebanyak 14 anggota. Nilai evaluasi SSE yang diperoleh adalah 2,4402 dan menghasilkan nilai Silhouette sebesar 0,5629. Penelitian lainnya yang dilakukan oleh (Toresa, 2020) dengan studi kasus “ Implementasi K-Means Terhadap Penyebaran Penyakit. TBC Di Riau Menggunakan Rapid Miner-”. Pada penelitiannya, dalam penentuan cluster terbaik peneliti tidak menggunakan pengukuran cluster optimum. Dari penelitian tersebut yaitu dengan menggunakan metode clustering K-Means pada data kasus Tuberculosis di Riau dengan wawancara langsung dengan BPS Riau Dan data tersebut di olah menggunakan aplikasi bantuan Rapid miner ditemukan dari 12 wilayah di Riau terdapat 3 daerah cluster tingkat rendah, 7 daerah cluster tingkat sedang dan 2 daerah memiliki cluster.

Dari permasalahan tersebut maka akan dilakukan penelitian dengan judul “Penerapan Algoritma *K-Means Clustering* Untuk Pengelompokan Penyebaran Penyakit TBC (Studi Kasus : Puskesmas di Kabupaten Jember)” untuk mengelompokkan puskesmas di kabupaten Jember berdasarkan data TBC pada tahun 2017 – 2019. Penelitian yang dilakukan yaitu menggunakan algoritma *K-Means* dengan metode *DBI (Davies Bouldin Index)* sebagai penentu cluster optimum dalam menentukan cluster terbaik. Penentuan cluster optimum akan diuji 2 sampai 10 cluster.

## 2. TINJAUAN PUSTAKA

### A. TBC

Tuberculosis (TBC) merupakan penyakit yang bisa menular dari granulomatosa yang sudah dikenal selama berabad-abad dan ditimbulkan oleh bakteri *Mycobacterium tuberculosis*. Di Indonesia wilayah yang paling terdampak kasus TBC berada pada wilayah Jawa timur pada Kabupaten Jember. Pada tahun 2017 sampai dengan 2019 orang yang terdampak penyakit TBC di Kabupaten Jember selalu meningkat. Berdasarkan hasil data dari Dinas Kesehatan Kabupaten Jember ada beberapa gejala dalam capaian kasus TBC yaitu:

1. Baru BTA + : (Pasien baru sudah cek lab Bakteri tahan asam positif),
2. BTA NEG ROG + : (Pasien sudah cek lab Bakteri tahan asam negatif *Rogensen* positif),
3. *Extra Paru* : (TBC yang diluar paru contohnya TB tulang, TB usus, TB saluran kemih dll.)
4. Kambuh - Gagal - DO dll. : (Sudah sembuh tapi kambuh, Minum obat 6 bulan di tengah jalan berhenti, Parah),
5. Kasus TBC Anak.

### B. Puskesmas

Pusat Kesehatan Masyarakat (Puskesmas) merupakan satu di antara yang ada fasilitas layanan kesehatan warga yang sangat berarti di Indonesia. Penafsiran puskesmas merupakan sesuatu unit pelaksana fungsional yang berperan selaku sentra pembangunan kesehatan, sentra pembinaan kedudukan dan warga di dalam bidang kebugaran dan pusat layanan kebugaran tingkatan awal dengan menyelenggarakan kegiatan tersebut secara merata, dan berkesinambungan terhadap warga yang bertempat tinggal di dalam sesuatu tempat tertentu (Azwar, 1996).

### C. Data Mining

*Data mining* merupakan sekumpulan data *knowledge* di dalam jumlah besar yang diolah membuat memperoleh suatu data, pola *knowledge* yang menarik dan juga sebagian pengetahuan. *Data mining* merupakan irisan berasal dari sebagian tekun ilmu yang lebih dahulu terdapat semacam statistika,

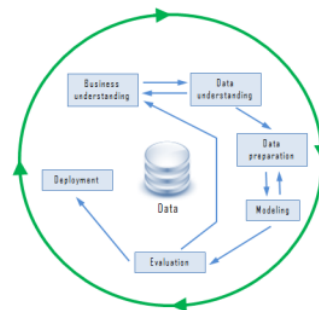
matematika, *artificial intelligence*, dan juga *machine learning* (Kamber & Pei, 2006).

### D. Clustering

Clustering adalah kegiatan mengelompokkan berbagai macam catatan pengetahuan, pengamatan, atau melakukan pendekatan pengetahuan dari satu obyek ke obyek lain tentunya menggunakan mempunyai kesamaan. Cluster merupakan deretan baris pengetahuan yang memiliki persamaan & perbedaan untuk dicatat ke pada cluster lain. Di mana kesamaan catatan pengetahuan meningkat sedangkan ketidaksamaan pengetahuan menurun muncul dari cluster kesamaan (Larose & Larose, 2014).

### E. Model CRISP DM

Proses *data mining* bersumber pada CRISP-DM berasal dari 6 fase. Yaitu:



Gambar 1 Fase CRISP DM  
Sumber : Indra Purnama Tahun 2014

2. **Business Understanding** adalah pemahaman tentang substansi dari kegiatan *data mining* yang akan dilakukan. Kegiatannya meliputi : pilih sasaran untuk menentukan tujuan bisnis, tahu keadaan bisnis, pilih tujuan *data mining* dan membawa dampak rencana strategi serta jadwal penelitian.
- b. **Data Understanding** artinya fase menghimpun data awal, menyelidiki data buat mampu mengenal data yang bisa dipakai, mengidentifikasi perkara yang berkenaan menggunakan kualitas data, mendeteksi sub set yang menarik berdasarkan data buat membawa efek hipotesa awal.



- c. **Data preparation** Aktivitas yang dilaksanakan dalam memilih table & field yang dapat dikonversi ke dalam basis data (dataset) baru buat **data mining** (set data mentah).
- d. **Modeling** (pemodelan) merupakan fase untuk menentukan teknik **data mining** yang digunakan, pilih **tools data mining**, teknik **data mining**, prosedur pemecahan **data mining**, pilih parameter menggunakan nilai optimal.
- e. **Evaluation** merupakan fase interpretasi terhadap **output data mining** yang ditunjukkan dalam proses - proses pemodelan terhadap fase sebelumnya. Evaluasi dilaksanakan secara mendalam menggunakan tujuan penyesuaian jenis yang diperoleh sehingga sinkronisasi menggunakan target yang diinginkan diperoleh fase pertama.
- f. **Deployment** atau penyebaran merupakan penyusunan laporan atau persiapan presentasi yang diperoleh sesuai dengan evaluasi eksplorasi data (Chapman, et al., 2000).

#### F. Algoritma K-means

Di antara banyak algoritma pengelompokan, algoritma pengelompokan K-means banyak digunakan karena algoritma yang sederhana dan konvergensinya yang cepat (Yuan & Yang, 2019). Tahapan perhitungan dasar Algoritma K-Means (Rahman, 2017) ialah:

1. Menentukan berapa banyak cluster yang ingin dibangun. Tetapkan pusat cluster
2. Memilah centroid secara acak sebanyak angka / jumlah cluster.
3. Menggunakan Euclidean distance dengan menghitung distance / jarak pada centroid, Gunakan rumus berikut ini :

Keterangan:

$$J(a_i b_i) = \sqrt{\sum_{i=1}^n (a_i b_i)^2}$$

$J(a_i b_i)$ : jarak objek antara nilai data dan centroid

- $a_i$  : nilai data dari dimensi
- $b_i$  : nilai jarak / centroid dari-dimensi
- $n$  : banyaknya-dimensi (atribut data)

4. Data dikelompokkan menurut jarak terdekat antara data dan centroid.
5. Gunakan rumus berikut untuk menentukan nilai centroid baru dan langkah-langkah untuk menghitung nilai rata-rata dari cluster terkait:

$$W_k = \frac{1}{n_k} \sum j_i$$

Keterangan:

- $n_k$  : jumlah data dalam cluster
- $j_i$  : Banyaknya nilai jarak yang terdapat pada setiap cluster

6. Ulangi langkah 3 sampai 5 hingga keanggotaan-setiap cluster tidak berubah atau tetap/sama.

#### G. Davies Bouldin Index (DBI)

Metode DBI ini pertama kali diusulkan oleh David L. Davies & Donald W. Bouldin pada tahun 1979. Evaluasi gunakan DBI ini memiliki skema penilaian evaluasi cluster internal, di mana **output cluster** diperoleh dari kuantitas & kedekatan antar data hasil cluster. Dalam penelitian (Widiarina & Wahono, 2015). Semakin terkecil nilai DBI menandakan skema cluster paling optimal. Rumus perhitungan Davies-Bouldin Index (DBI) sebagai berikut (Al-Anazi, AlMahmoud, & Al-Turaiki, 2016) :  
 Di mana :

- DB : Davies Bouldin
- $n$  : jumlah cluster
- $\sigma_i$  : rata-rata jarak dari data dengan titik pusat data cluster ke-i

$$DB = \frac{1}{n} \sum_{i=1, i \neq j}^n \max \left( \frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right)$$

- $\sigma_j$  : rata-rata jarak dari data dengan titik pusat data cluster ke-j
- $c_i$  : titik pusat data cluster i
- $c_j$  : titik pusat data cluster j
- $d(c_i, c_j)$  : jarak antara centroid atau medoid dari  $c_i$  dan  $c_j$

#### H. Rapidminer Studio

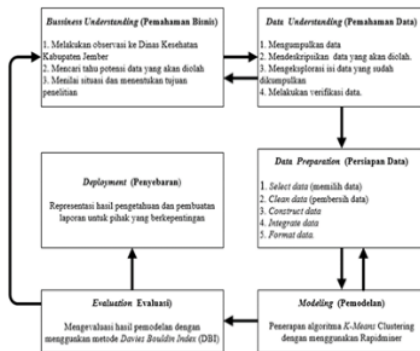
RapidMiner adalah perangkat lunak atau perangkat lunak open source (sumber terbuka), sebagai mesin data mining, dapat digunakan secara independen untuk kesimpulan atau

analisis data, dan dapat diintegrasikan ke dalam produknya sendiri. RapidMiner dapat merancang saluran analisis dengan menyediakan GUI (antarmuka pengguna grafis) yang menghasilkan file XML (Extensible Markup Language). RapidMiner kemudian membaca file ini untuk menjalankan analisis secara otomatis (Prmono, Saputra, SA, & Burhanuddin, 2018).

### 3. METODOLOGI PENELITIAN

#### A. Cross-industry standard process model for data mining (CRISP-DM).

CRISP DM membawa 6 tahapan investigasi, yakni (business understanding) pemahaman bisnis, (data understanding) pemahaman data, (data ready) persiapan data, (modeling) pemodelan, (evaluation) penilaian, dan penyebaran (deployment). Eksplorasi atau penelitian ini melakukan metode algoritma K-means clustering untuk pengelompokan data persoalan TBC di Kabupaten Jember pada tahun 2017 – 2019 tanpa pembuatan sebuah program. Informasi atau data tersebut diperoleh bersumber berdasarkan dari Dinas Kesehatan Kabupaten Jember. Untuk pemilihan jumlah cluster yang akan di bentuk menggunakan metode DBI. Dataset yang telah diperoleh akan di olah menggunakan perangkat / tools Rapidminer.



Gambar 2. Tahapan Penelitian. dengan CRISP-DM

Sumber : Toresa Tahun 2020

#### a. Business Understanding (Pemahaman Bisnis)

Tahap pemahaman- bisnis atau sering disebut Business Understanding adalah cara untuk memilih alasan eksplorasi. Tujuan penelitian ini adalah untuk mengelompokkan daerah penyebaran TBC di Puskesmas di wilayah Jember dengan menggunakan perhitungan K-Means. Latihan persepsi atau kegiatan observasi dilakukannya secara eksklusif / pribadi kepada-Dinas Kesehatan (DINKES)-Kabupaten Jember untuk mengetahui kapasitas data yang akan disiapkan maupun diolah. Hasil pemeriksaannya, diharapkan dapat membantu Pemerintah-Kabupaten Jember untuk menyelesaikan upaya penanganan infeksi penyakit TBC.

#### b. Data Understanding (Pemahaman Data)

Tahap selanjutnya ialah pemahaman data atau data understanding. Dataset dalam penelitian ini adalah data tentang masalah TBC di Puskesmas Kabupaten Jember pada tahun 2017-2019.

Berikut adalah tabel hasil yang mendeskripsikan data dan mengeksplorasi data:

Tabel 1. Deskripsi Data Perkara TBC

Atribut	Tipe Data	Keterangan
NO	Numerik	Nomor urutan daftar Puskesmas
PUSKESMAS	Varchar	Puskesmas di Kabupaten Jember
BARU BTA +	Numerik	Pasien baru sudah cek lab Bakteri tahan asam positif
BTA NEG ROG+	Numerik	Pasien sudah cek lab Bakteri tahan asam negatif Rogosen positif
EP	Numerik	TBC yang diluar paru contohnya : TBC tulang belakang, TBC usus, TBC saluran kencing dan alat kelamin
KAMBUH GAGAL DO	Numerik	Sudah sembuh tapi kambuh , Minum obat 6 bulan ditengah jalan berhenti , Parah
KASUS ANAK	Numerik	Total Kasus TBC anak di suatu Puskesmas

Sumber : Penelitian Sendiri

Tabel 1 merupakan deskripsi data perkara TBC pada Puskesmas Kabupaten Jember dalam tahun 2017-2019 . Menggambarkan informasi / data yang dapat dipertanggungjawabkan untuk menggambarkan atribut, jenis data, dan kenyataan tentang gambaran karakteristik data.

Tabel 2 Ekspoler Data Perkara TBC

Atribut	Tipe Data	Keterangan
NO	Numerik	Nomor urutan daftar Puskesmas pada Tahun 2017-2019 : 1-50 Tahun 2017-2019 : Terdapat 50 Puskesmas
PUSKESMAS	Varchar	Nama Puskesmas yang terdapat di Kabupaten Jember Tahun 2017-2019 : Terdapat 50 Puskesmas
BARU BTA +	Numerik	Pasien baru sudah cek lab Bakteri tahan asam positif Tahun 2017: 2285 Pasien Tahun 2018: 2335 Pasien Tahun 2019: 2643 Pasien
BTA NEG ROG+	Numerik	Pasien sudah cek lab Bakteri tahan asam negatif Rogensen positif Tahun 2017: 632 Pasien Tahun 2018: 546 Pasien Tahun 2019: 731 Pasien
EP	Numerik	TBC yang ditular paru contohnya TBC tulang belakang, TBC usus, TBC saluran kencing dan alat kelamin Tahun 2017: 272 Pasien Tahun 2018: 310 Pasien Tahun 2019: 394 Pasien
KAMBUH GAGAL DO	Numerik	Sudah sembuh tapi kambuh, Minum obat 6 bulan ditengah jalan berhenti, Parah Tahun 2017: 184 Pasien Tahun 2018: 222 Pasien Tahun 2019: 216 Pasien
KASUS ANAK	Numerik	Total Kasus TBC anak di suatu Puskesmas Tahun 2017: 212 Kasus Tahun 2018: 206 Kasus Tahun 2019: 273 Kasus

Sumber : Penelitian Sendiri

Tabel 2 Menggambarkan menyelidiki data yang tugasnya memperoleh data tentang atribut atau kualitas informasi dan data yang digunakan, dengan tujuan agar data tersebut lebih jelas.

**c. Data Preparation (Persiapan Data)**

Data Preparation disebut juga (Persiapan data) adalah seluruh latihan / aktivitas penanganan data yang mendasari yang dimaksudkan untuk memberikan dataset terakhir / final. Dalam penelitian ini, data kasus TBC tidak perlu di bersihkan, Dilihat dari datanya maka akan dilakukan proses selanjutnya yaitu pemodelan karena tidak ada data yang missing value dan tidak perlu dinormalisasi sebab tipe data numerik Baru BTA + tidak berjauhan dari tipe lainnya

Tabel 3. TBC Setelah Cleaning

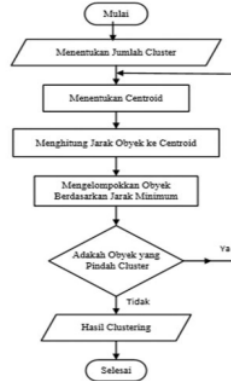
NO	PUSKESMAS DI KABUPATEN JEMBER	JUMLAH KASUS TB TAHUN 2019 TW 1 sd TW 4				
		BARU BTA +	BTA NEG RO+	EP	KAMBU H-GAGAL	KAT ANAK
1	PUSKESMAS KENCONG	46	10	8	6	6
2	PUSKESMAS CAKRU	19	1	1	1	2
3	PUSKESMAS PUGER	23	47	10	10	5
4	PUSKESMAS KASİYAN	69	4	8	1	3
5	PUSKESMAS LOJEJER	47	0	4	2	1
6	PUSKESMAS AMBULU	36	10	12	5	1
7	PUSKESMAS SILO II	39	20	11	3	5
8	PUSKESMAS MAYANG	60	7	5	0	6
9	PUSKESMAS BALUNG	57	6	8	4	3
10	PUSKESMAS SUKOREJO	29	18	6	2	2
11	PUSKESMAS BANGSALSARI	71	20	10	3	5
12	PUSKESMAS SUTOWONO	50	13	13	6	7
13	PUSKESMAS ELBUE	31	5	2	6	8
14	PUSKESMAS PATRANG	12	98	1	6	12
15	PUSKESMAS BANJARSENGON	39	0	2	3	9

Keterangan : TW = Tri Wulan

Sumber : Penelitian Sendiri

**d. Modeling (Pemodelan).**

Langkah-langkah perhitungannya



Gambar 3. Flowchart Alur Algoritma K-Means

Sumber : Rahman Tahun 2017

Sampel yang diambil dari data-Penyakit TBC di Puskesmas Kabupaten Jember pada tahun 2017 – 2019. Yang akan diuji cobakan data sampel terdiri dari 15 puskesmas pada tahun 2019. proses algoritma K-Means dalam pengelompokan Puskesmas berdasarkan penyakit. TBC di Kabupaten Jember adalah sebagai berikut:

Tabel 4. Pengelompokan Data Cluster Iterasi 1

NO	PUSKESMAS DI KABUPATEN JEMBER	C1 (TBC)	C2 (TBC)	C1	C2
1	PUSKESMAS KENCONG	12,569805	94,78924	OK	
2	PUSKESMAS CAKRU	28,213472	97,892799	OK	
3	PUSKESMAS PUGER	53,860932	53,553711		OK
4	PUSKESMAS KASİYAN	22,825424	110,63453	OK	
5	PUSKESMAS LOJEJER	0	104,76163	OK	
6	PUSKESMAS AMBULU	17,146428	92,536479	OK	
7	PUSKESMAS SILO II	23,021729	83,492515	OK	
8	PUSKESMAS MAYANG	15,748016	103,31021	OK	
9	PUSKESMAS BALUNG	12,649111	103,06794	OK	
10	PUSKESMAS SUKOREJO	25,553865	82,643814	OK	
11	PUSKESMAS BANGSALSARI	32,07803	98,508883	OK	
12	PUSKESMAS SUKOWONO	17,635192	94,010638	OK	
13	PUSKESMAS JELBUK	18,708287	95,010526	OK	
14	PUSKESMAS PATRANG	104,76163	0	OK	OK
15	PUSKESMAS BANJARSENGON	11,532563	101,74478	OK	
TOTAL				13	2

Sumber : Penelitian Sendiri

Tabel 4 adalah hasil pengelompokan. Perhitungan dapat bisa dicoba secara manual dengan memakai microsoft excel. Nilai minimum adalah nilai pilihan jika nilai terkecil (minimum) dapat dimasukkan dalam cluster. Dalam penelitian ini, data dihitung ulang sampai iterasi ke 2, di mana setiap cluster tidak

bergerak dari satu *cluster* ke *cluster* lainnya atau sama.

Tabel 5 Centroid Akhir

Centroid Awal (Items 1)	Data ke -5 Sebagai Cluster 1	Data ke -39 Sebagai Cluster 2
BARU BTA +	45,61538462	17,5
BTA NEG ROG +	8,769230769	72,5
EP	6,923076923	5,5
KAMBUH-GAGAL-DO-DLL	3,230769231	8
KAT ANAK	4,461538462	8,5

Sumber : Penelitian Sendiri

**e. Evaluation (Evaluasi)**

Pada tahap selanjutnya yaitu evaluasi atau penilaian, hasil pengelompokan yang diperoleh dari jumlah cluster  $K = 2$  dinilai menggunakan (Davies Bouldin Index) DBI metode validasi untuk menguji sebuah cluster yang paling baik. Nilai terbesar yang dipilih untuk dicari rata-ratanya menghasilkan nilai DBI sebesar 0,608

Tabel 6. Davies-Bouldin Indeks

$\frac{\sigma_1 + \sigma_2}{d(C_1, C_2)}$	Cluster 1	Cluster 2	MAX	DBI
1	0	0,607506704	0,607506704	0,608
2	0,607506704	0	0,607506704	

Sumber : Penelitian Sendiri

**f. Deployment (Penyebaran)**

Pada tahap akhir yaitu Penyebaran (Deployment) adalah penggambaran informasi atau data bersumber dari data yang sudah disusun dalam pengukuran proses data mining. Tujuan peruntukan dalam eksplorasi atau pemeriksaan ini adalah akomodasi output / hasil eksplorasi riset ke (DINKES) Dinas-Kesehatan Kabupaten-Jember sebagai penilaian atau mengevaluasi untuk menangani kasus perkara penyakit TBC, sehingga diyakini kuantitas keseluruhan permasalahan TBC di wilayah Kab. Jember sanggup menurun atau berkurang.

**4. HASIL DAN PEMBAHASAN**

**A. Data Pengujian**

Data yang di uji data dari Dinas Kesehatan Kabupaten Jember. Yaitu data penyakit TBC terdapat 50 puskesmas di Kabupaten Jember pada tahun 2017 sampai dengan 2018 dengan 5 Atribut yaitu Baru BTA +, BTA NEG ROG+,

Extra Paru, Kambuh -gagal -DO dll, dan Kategori Anak.

**B. K-means pada Rapid Miner**

Proses *clustering* algoritma *K-Means* dapat dihitung menggunakan *tools RapidMiner*.



Gambar 4. Proses *Evaluasi* kinerja *K-Means*  
 Sumber : Penelitian Sendiri

**C. Penentuan Jumlah Cluster Optimum**

*Cluster* terbaik didapatkan dari hasil perhitungan cluster dengan nilai DBI terkecil dalam 3 tahun.

Tabel 7. Nilai DBI Tahun 2017, 2018 dan 2019

Cluster ke	Cluster yang di hasilkan	Nilai DBI	Anggota Cluster
2	2	0,833	Cluster 0: 45 items, dan Cluster 1: 5 items.
3	3	0,998	Cluster 0: 27 items, Cluster 1: 21 items, dan Cluster 2: 2 items.
4	4	0,897	Cluster 0: 23 items, Cluster 1: 21 items, Cluster 2: 2 items dan Cluster 3: 4 items.
5	5	1,044	Cluster 0: 15 items, Cluster 1: 2 items, Cluster 2: 22 items Cluster 3: 4 items, dan Cluster 4: 7 items.
6	6	1,059	Cluster 0: 14 items, Cluster 1: 13 items, Cluster 2: 1 items Cluster 3: 14 items, Cluster 4: 4 items, dan Cluster 5: 4 items.
7	7	0,977	Cluster 0: 13 items, Cluster 1: 4 items, Cluster 2: 1 items, Cluster 3: 1 items, Cluster 4: 14 items, Cluster 5: 4 items, dan Cluster 6: 13 items.
8	8	0,885	Cluster 0: 3 items, Cluster 1: 13 items, Cluster 2: 4 items, Cluster 3: 14 items, Cluster 4: 1 items, Cluster 5: 1 items, Cluster 6: 1 items, dan Cluster 7: 13 items
9	9	0,784	Cluster 0: 1 items, Cluster 1: 12 items, Cluster 2: 1 items, Cluster 3: 1 items, Cluster 4: 1 items, Cluster 5: 4 items, Cluster 6: 13 items, Cluster 7: 3 items, dan Cluster 8: 14 items
10	10	0,803	Cluster 0: 4 items, Cluster 1: 1 items, Cluster 2: 13 items, Cluster 3: 1 items, Cluster 4: 12 items, Cluster 5: 1 items, Cluster 6: 12 items, Cluster 7: 3 items, Cluster 8: 1 items, dan Cluster 9: 2 items

Sumber : Penelitian Sendiri

Tabel 7. menunjukkan nilai Davies Bouldin Indeks pengujian 2 s/d 10 cluster, nilai yang paling terkecil yaitu berada di 9 cluster dengan nilai Davies Bouldin Indeks yaitu 0,784.



**D. Hasil Profiling Cluster**

Didapatkan jumlah *cluster* optimal dan karakteristik *cluster* yang berbeda, berikut ini merupakan penjelasan karakteristik anggota puskesmas yang terbentuk.

**Tabel 8. Anggota Cluster 1 Kasus TBC**

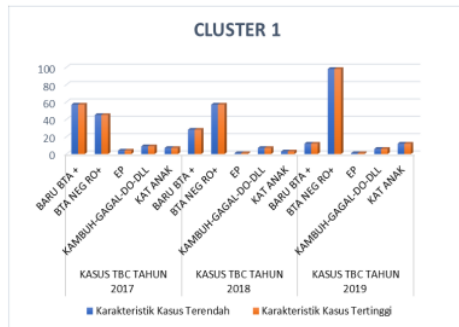
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 1														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019								
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KA MB UH GA GAL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
49	PATRANG	57	45	4	9	7	28	57	1	7	3	12	98	1	6	12

Sumber : Penelitian Sendiri

**Tabel 9. Karakteristik Cluster 1 Kasus TBC**

Cluster	Jumlah Anggota	CLUSTER 1														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019								
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KA MB UH GA GAL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
		57	45	4	9	7	28	57	1	7	3	12	98	1	6	12
		57	45	4	9	7	28	57	1	7	3	12	98	1	6	12

Sumber : Penelitian Sendiri



**Gambar 5. Diagram Karakteristik Cluster 1**

Sumber : Penelitian Sendiri

**Tabel 10. Anggota Cluster 2 Kasus TBC**

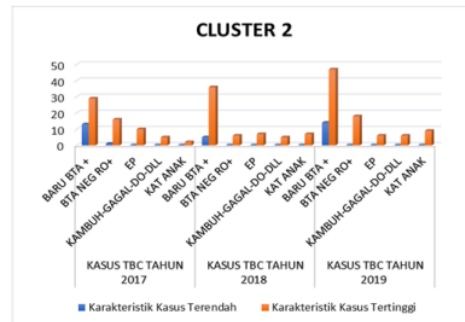
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 2														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019								
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KA MB UH GA GAL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
2	CAKRU	22	1	1	0	0	34	0	2	1	2	19	1	1	1	2
4	TEMBOKREJO	18	1	5	1	0	22	1	0	1	1	27	1	0	0	1
8	LOJER	17	16	3	2	0	31	3	2	0	0	47	0	4	2	1
10	SABRANG	19	10	5	5	2	22	3	9	4	2	14	7	3	3	4
13	CURAHJONGGKO	28	1	0	1	2	27	0	0	0	0	28	0	0	2	0
22	NOGOSARI	15	8	6	1	1	29	6	3	3	1	22	11	6	5	2
32	SUKOREJO	29	1	0	1	0	31	2	1	3	3	29	18	6	2	2
34	KLATAKAN	17	7	1	0	2	5	6	7	0	0	21	11	5	0	4
43	REBUK	20	15	2	2	1	28	3	4	5	7	31	5	2	6	8
45	MANGLI	18	9	10	1	1	36	4	7	2	5	25	3	4	6	6
48	GLADAKPAKEM	13	11	3	3	1	9	5	6	1	1	23	0	7	2	2
50	BANGSARENGON	22	3	7	1	1	36	1	3	5	2	39	0	2	3	9

Sumber : Penelitian Sendiri

**Tabel 11. Karakteristik Cluster 2 Kasus TBC**

Cluster	Jumlah Anggota	CLUSTER 2														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019								
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KA MB UH GA GAL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
		13	1	0	0	0	5	0	0	0	0	14	0	0	0	0
		29	16	10	5	2	36	6	7	5	7	47	18	6	6	9

Sumber : Penelitian Sendiri



**Gambar 6. Diagram Karakteristik Cluster 2**

Sumber : Penelitian Sendiri

**Tabel 12. Anggota Cluster 3 Kasus TBC**

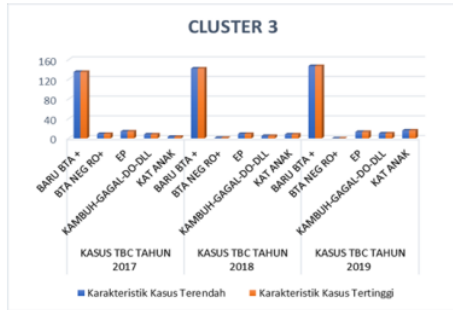
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 3														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019								
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KA MB UH GA GAL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
47	SUMBERSARI	135	9	14	8	3	142	1	9	5	8	147	0	13	10	16

Sumber : Penelitian Sendiri

**Tabel 13. Karakteristik Cluster 3 Kasus TBC**

Cluster	Jumlah Anggota	CLUSTER 3														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019								
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KA MB UH GA GAL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
		135	9	14	8	3	142	1	9	5	8	147	0	13	10	16
		135	9	14	8	3	142	1	9	5	8	147	0	13	10	16

Sumber : Penelitian Sendiri



Gambar 7. Diagram Karakteristik Cluster 3  
 Sumber : Penelitian Sendiri

Tabel 14. Anggota Cluster 4 Kasus TBC

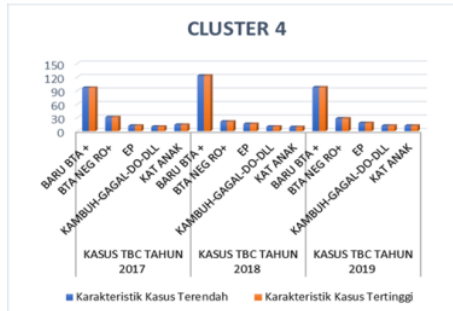
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 4														
		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019						
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
39	KALISAT	96	31	12	10	14	123	21	16	10	9	97	28	18	12	12

Sumber : Penelitian Sendiri

Tabel 15. Karakteristik Cluster 4 Kasus TBC

Cluster	Jumlah Anggota	CLUSTER 4															
		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019							
Karakteristik	CLUSTER 4	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	
		Kasus Terendah	96	31	12	10	14	123	21	16	10	9	97	28	18	12	12
		Kasus Tertinggi	96	31	12	10	14	123	21	16	10	9	97	28	18	12	12

Sumber : Penelitian Sendiri



Gambar 8. Diagram Karakteristik Cluster 4  
 Sumber : Penelitian Sendiri

Tabel 16. Anggota Cluster 5 Kasus TBC

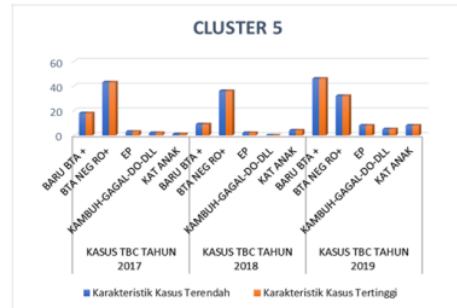
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 5														
		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019						
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
29	SUMBERBARU	18	43	3	2	1	9	36	2	0	4	46	32	8	5	8

Sumber : Penelitian Sendiri

Tabel 17. Karakteristik Cluster 5 Kasus TBC

Cluster	Jumlah Anggota	CLUSTER 5															
		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019							
Karakteristik	CLUSTER 5	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	
		Kasus Terendah	18	43	3	2	1	9	36	2	0	4	46	32	8	5	8
		Kasus Tertinggi	18	43	3	2	1	9	36	2	0	4	46	32	8	5	8

Sumber : Penelitian Sendiri



Gambar 9. Diagram Karakteristik Cluster 5  
 Sumber : Penelitian Sendiri

Tabel 18. Anggota Cluster 6 Kasus TBC

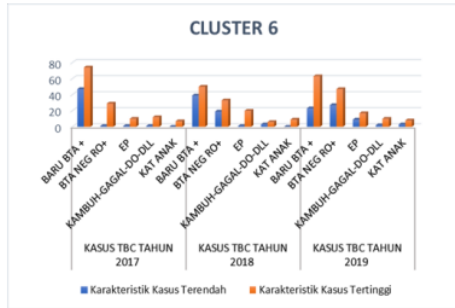
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 6														
		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019						
		BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K
3	GUMUKMAS	61	22	2	3	0	48	20	1	5	0	26	48	9	2	4
5	PUGOER	47	29	10	12	2	40	25	8	5	3	23	47	10	10	5
21	RAMBIPUJI	74	20	6	4	7	50	19	13	6	7	50	27	17	7	3
31	TANGGUL	58	1	1	1	1	39	20	3	9	63	37	10	4	8	

Sumber : Penelitian Sendiri

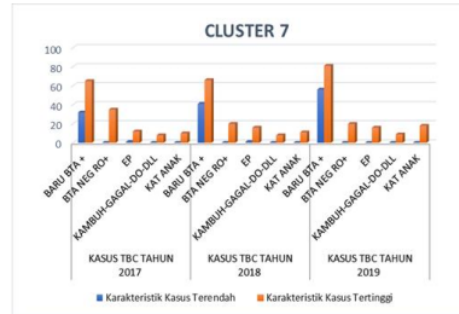
Tabel 19. Karakteristik Cluster 6 Kasus TBC

Cluster	Jumlah Anggota	CLUSTER 6															
		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019							
Karakteristik	CLUSTER 6	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	BAR U BTA +	BTA NEG RO+	EP	KAM BUH GAG AL-DO-DLL	KAT ANA K	
		Kasus Terendah	47	1	1	1	0	39	19	1	3	0	23	27	9	2	3
		Kasus Tertinggi	74	20	10	12	7	50	33	20	6	9	63	47	17	10	8

Sumber : Penelitian Sendiri



Gambar 10. Diagram Karakteristik Cluster 6  
 Sumber : Penelitian Sendiri



Gambar 11. Diagram Karakteristik Cluster 7  
 Sumber : Penelitian Sendiri

Tabel 20. Anggota Cluster 7 Kasus TBC

NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 7														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019			KAT ANAK					
		BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK
6	KASIHAN	42	2	1	2	2	52	3	6	0	1	69	4	8	1	3
16	MAYANG	63	13	3	2	1	64	7	1	3	1	60	7	5	0	6
17	MUMBELARI	60	35	9	3	2	49	20	16	1	6	69	16	6	1	4
18	JENGGAWAH	53	11	6	6	0	46	8	8	3	0	77	19	10	3	6
19	KEMUNINGSAR	51	9	3	2	0	50	4	4	1	0	56	7	8	4	4
23	BALUNG	47	3	3	4	1	41	4	6	4	1	57	6	8	4	3
28	JOMBANG	53	1	3	3	1	46	0	1	3	1	63	2	6	7	0
33	BANGSALSARI	65	17	6	6	10	63	8	8	4	5	71	20	10	3	5
36	SUKORAMBI	44	0	6	2	1	59	0	4	4	11	81	0	4	0	15
38	PAKUSARI	63	16	2	3	3	66	3	3	3	2	66	10	5	3	3
40	LEDOKOMBO	60	9	11	0	5	49	11	3	1	2	76	14	0	0	1
44	KALIWATES	52	10	11	3	5	51	15	3	7	10	67	8	11	7	18
46	JEMBERKIDUL	32	23	12	8	2	63	4	11	8	3	68	11	16	9	10

Sumber : Penelitian Sendiri

Tabel 21. Karakteristik Cluster 7 Kasus TBC

Karakteristik	Cluster 7	KASUS TBC TAHUN														
		2017			2018			2019			KAT ANAK					
		BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK
	Kasus Terendah	32	0	1	0	0	41	0	1	0	0	56	0	0	0	0
	Kasus Tertinggi	65	35	12	8	10	66	20	16	8	11	81	20	16	9	18

Sumber : Penelitian Sendiri

Tabel 22. Anggota Cluster 8 Kasus TBC

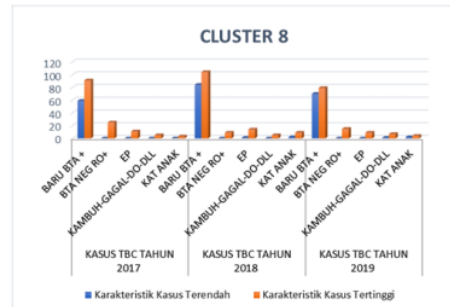
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 8														
		KASUS TBC TAHUN 2017			KASUS TBC TAHUN 2018			KASUS TBC TAHUN 2019			KAT ANAK					
		BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK
20	AJUNG	59	25	11	5	2	91	9	14	5	2	79	15	9	3	2
25	UMBELSARI	71	9	3	3	0	104	0	1	1	9	70	0	8	1	4
35	PANTI	91	0	0	0	3	84	1	1	0	5	72	8	0	7	2

Sumber : Penelitian Sendiri

Tabel 23. Karakteristik Cluster 8 Kasus TBC

Karakteristik	Cluster 8	KASUS TBC TAHUN														
		2017			2018			2019			KAT ANAK					
		BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK	BAR U BTA +	BTA NEG RO+	EP	KAM BUIH-GAGAL-DO-DOL	KAT ANAK
	Kasus Terendah	59	0	0	0	0	84	0	1	0	2	70	0	0	1	2
	Kasus Tertinggi	91	25	11	5	3	104	9	14	5	9	79	15	9	7	4

Sumber : Penelitian Sendiri



Gambar 12. Diagram Karakteristik Cluster 8  
 Sumber : Penelitian Sendiri

Tabel 24. Anggota Cluster 9 Kasus TBC

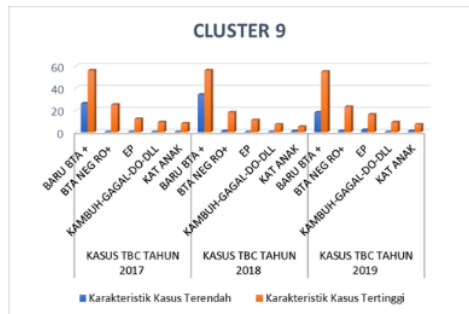
NO	PUSKESMAS DI KABUPATEN JEMBER	CLUSTER 9														
		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019						
		BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K	BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K	BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K			
1	KENCONG	51	15	8	6	3	49	8	7	7	3	46	10	8	6	6
7	WULUHAN	44	9	7	1	4	56	10	3	3	3	45	16	11	5	7
9	AMBULU	44	7	7	4	8	56	6	11	5	1	36	10	12	5	1
11	ANDONGSARI	39	5	6	0	0	35	10	5	0	1	18	11	8	0	1
12	TEMPURJO	56	0	0	4	4	34	18	3	1	1	40	18	7	2	5
14	SILU I	39	11	7	2	4	39	9	6	8	5	31	5	4	3	3
15	SILU II	37	22	10	2	2	48	13	8	5	5	39	20	11	3	5
24	KARANGDUREN	38	4	4	5	2	45	9	10	2	3	40	16	5	4	1
26	PALERAN	45	0	1	2	0	49	1	0	3	2	35	1	2	3	1
27	SEMBORO	29	7	1	4	0	39	7	0	3	1	52	5	5	9	4
30	ROWOTENGAH	27	4	2	3	3	45	5	3	3	5	55	11	7	1	5
37	ARJASA	40	19	12	8	8	41	6	6	5	5	40	6	16	7	5
41	SUMBERHAMBU	26	25	9	3	2	46	11	9	4	2	29	23	10	7	6
42	SUKOWONO	52	15	7	9	1	41	9	5	5	2	50	13	13	6	7

Sumber : Penelitian Sendiri

Tabel 25. Karakteristik Cluster 9 Kasus TBC

Cluster		CLUSTER 9														
Jumlah Anggota		KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019						
Karakteristik		BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K	BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K	BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K			
		Kasus Terendah	26	0	0	0	34	1	0	0	1	18	1	2	0	1
		Kasus Tertinggi	56	25	12	9	8	56	18	11	7	5	55	23	16	9

Sumber : Penelitian Sendiri



Gambar 13. Diagram Karakteristik Cluster 9  
 Sumber : Penelitian Sendiri

Berdasarkan Tabel dan diagram maka akan dilakukan *profiling cluster*, berikut ini penjelasan karakteristik cluster yang terbentuk pada setiap atribut.

1. Baru BTA + pada cluster 3 di tahun 2017 sampai 2019 memiliki nilai lebih tinggi dari pada cluster lainnya.

2. BTA NEG ROG + di tahun 2017 sampai 2019, cluster 1 mendapatkan nilai yang lebih tinggi dibanding cluster lainnya.
3. Extra Paru pada cluster 6 ditahun 2017 sampai 2019 unggul lebih tinggi di dibandingkan cluster lainnya.
4. Kambuh-Gagal-DO dll yang paling tertinggi berada pada cluster 4 sama dibandingkan cluster lainnya.
5. Terakhir, cluster 7 mendapatkan nilai tertinggi dibandingkan cluster lainnya untuk Kategori Anak di tahun 2017 sampai 2019.

Selanjutnya untuk mengetahui anggota cluster mana yang menempati nilai tinggi dan rendah maka akan dilakukan dengan membandingkan rata-rata data awal dengan rata-rata dari setiap cluster yang dihasilkan.

Jika :

1. Nilai rata - rata data awal > ( Lebih dari ) nilai rata - rata cluster yang dihasilkan maka, Hasilnya tinggi. Dan
2. Nilai rata - rata data awal < ( Kurang dari ) nilai rata - rata cluster yang dihasilkan maka, Hasilnya Rendah (Zakiyyah, Rosyidah, & Margaretha, 2021).

Tabel 26. Hasil karakteristik di setiap cluster

Cluster & tahun	KASUS TBC TAHUN 2017				KASUS TBC TAHUN 2018				KASUS TBC TAHUN 2019				Kategori			
	BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K	BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K	BARU BTA +	BTA NEG RO+	KAM BUB GAGA L-DO-DLL	KAT ANA K	Tinggi	Rendah		
Cluster 1	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	9	6		
Cluster 2	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	-	15		
Cluster 3	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	12	3		
Cluster 4	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	15	-		
Cluster 5	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	7	8		
Cluster 6	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	10	5		
Cluster 7	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	8	7		
Cluster 8	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	4	11		
Cluster 9	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	Tinggi	6	9		
Tinggi	6	4	4	5	5	4	4	4	5	5	4	4	6	5	71	
Rendah	3	5	5	4	4	4	5	5	4	4	5	5	3	3	4	44

Sumber : Penelitian Sendiri

Dari cluster 1 - 9 untuk atribut Baru BTA + tahun 2017 memiliki nilai lebih tinggi di bandingkan tahun 2018 dan 2019. Untuk atribut BTA NEG ROG + tahun 2017 s/d 2019 memiliki nilai yang sama, pada atribut Extra Paru tahun 2017 dan 2018 memiliki nilai lebih rendah dibandingkan tahun 2019, untuk atribut Kambuh - Gagal - DO dll tahun 2017 dan 2018 memiliki nilai sama rendah sedangkan di tahun 2019 memiliki nilai tinggi, dan pada atribut



Kategori anak ditahun 2017 sampai 2018 memiliki nilai sama rendah di bandingkan tahun 2019 memiliki nilai tinggi. Menunjukkan bahwa kategori nilai kasus tingkat tinggi selama 3 tahun di semua atribut berada di cluster 4 dan nilai kasus rendah berada pada cluster 2.

## 5. KESIMPULAN DAN SARAN

### A. KESIMPULAN

Berdasarkan hasil penelitian yaitu pengujian dan analisa profiling cluster dengan menggunakan algoritma K-means clustering dan uji validasi Davies Bouldin Index, maka didapatkan kesimpulan bahwa :

1. Hasil dari penerapan metode *K-means clustering* dalam mengelompokkan puskesmas di Kabupaten Jember pada kasus penyakit TBC diperoleh *cluster* optimum dengan hasil 9 *cluster* berdasarkan *Davies Bouldin Index* dengan nilai sebesar 0,784 dari skenario *cluster* 2 sampai dengan 10 *cluster*.
2. Hasil anggota masing - masing cluster pada cluster optimum yaitu cluster 1 terdapat 1 puskesmas, pada cluster 2 terdapat 12 puskesmas, untuk cluster 3, 4, 5 terdapat 1 masing - masing puskesmas, pada cluster 6 terdapat 4 puskesmas, pada cluster 7 terdapat 13 puskesmas, pada cluster 8 terdapat 3 puskesmas, dan cluster 9 terdapat 14 puskesmas.
3. Berdasarkan karakteristik cluster optimum menunjukkan bahwa kategori nilai kasus yang tertinggi pada tahun 2017 sampai 2019 berada pada Baru BTA + di cluster 3, BTA NEG ROG + di cluster 1, Extra Paru berada di cluster 6, Kambuh Gagal - Do - dll berada di cluster 4 dan kategori anak menempati di cluster 7. Untuk kategori nilai kasus tingkat tinggi selama 3 tahun di semua atribut berada pada cluster 4 sedangkan nilai kasus tingkat rendah berada pada cluster 2.

### B. SARAN

Berdasarkan penelitian yang telah dilakukan maka didapatkan saran bahwa:

1. Penelitian ini hanya sebagai penerapan dari algoritma K-Means Clustering, agar lebih bermanfaat maka disarankan membangun

Sistem Informasi sebagai media implementasi algoritma K-Means.

2. Penelitian ini dapat dikembangkan menggunakan metode clustering yang berbeda, untuk uji validasi cluster dapat menggunakan teknik yang berbeda juga Dan dapat meningkatkan jumlah atribut dan variabel pada data yang digunakan.

## 6. REFERENSI

- Al-Anazi, S., AlMahmoud, H., & Al-Turaiki, I. (2016). Finding Similar Documents Using Different Clustering Techniques. *Procedia Computer Science*, 82, 28-34.
- Azwar, A. (1996). Manajemen pelayanan kesehatan. Jakarta : Banacipta.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide.SPSS inc., 9, 13.
- Dhuhita, W. M. (2005). Clustering Menggunakan Metode K-Means untuk Menentukan Status Gizi Balita. *Jurnal Informatika*, 15(2), 160-174.
- Fadillah, A. (2015). Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ). *JuTISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 1(3).
- Fauzi, M. (2018). Penerapan Algoritma K-Means Clustering Untuk Mendeteksi Penyebaran Penyakit TBC (Studi Kasus: Di Kabupaten Deli Serdang). *JTIK (Jurnal Teknik Informatika Kaputama)*, 1(2), 1-7.
- Indah, M. (2018). "Infodatin, Pusat Data Informasi Kemenrian Kesehatan RI". Retrieved januari 2019., from <https://pusdatin.kemkes.go.id/resources/download/pusdatin/infodatin/infodatin-tuberkulosis-2018.pdf>.
- Irwan Budiman, M., Prahasto, I. T., ASc, M., & Yuli Christiyono, S. T. (2012). Data Clustering menggunakan metodologi Crisp-DM untuk pengenalan pola proporsi pelaksanaan tridharma. (Doctoral dissertation, Universitas Diponegoro).

- Kader, J. A., Napitapulu, D. B., & Jati, R. K. (2017). Clustering SMEs Based on Technology Readiness using K-Means Algorithm. *Jurnal Penelitian Pos dan Informatika*, 7(2), 97-108.
- Kamber, M., & Pei, J. (2006). *Data Mining. Morgan kaufmann*.
- Larose, D., & Larose, C. (2014). *Discovering Knowledge in Data : An Introduction* (Vol. 4). John Wiley & Sons.).
- Ong, J. O. (2013). 'Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university. *Jurnal Ilmiah teknik Industri*, 12 (1), 10–20.
- Parasuraman, A., & Colby, C. L. (2015). An Updated and Streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research*, 18(1), 59–74. Retrieved from <https://doi.org/10.1177/1094670514539730>
- Prajitno, A. A., & Purwarianti, A. (2012). Prediksi Kinerja Penjualan Karya Musik Menggunakan Framework CRISP-DM (Studi Kasus: X Music Indonesia). *Jurnal Sarjana ITB bidang Teknik Elektro dan Informatika*, 1(1).
- Pramono, F., Saputra, SA, & Burhanuddin, K. (2018). Komparasi Klasifikasi Penentuan Keterlambatan Siswa SMA Datang Menggunakan Algoritma C4.5. *Dalam Seminar Nasional Teknologi Informasi dan Komunikasi (Sentika)*, Halaman (hlm. 80-86).
- Purnama, I., Saputra, R., & Wibowo, A. (2014). Implementasi Data Mining Menggunakan Crisp-Dm Pada Sistem Informasi Eksekutif Dinas Kelautan Dan Perikanan Provinsi Jawa Tengah. *Dalam Prosiding Seminar Nasional Ilmu Komputer UNDIP 2012*.
- Rahman, A. T. (2017). Pengelompokan data perdagangan batubara menggunakan K-means (studi kasus Pt. Global Bangkit Utama). *ITSMART: Jurnal Teknologi dan Informasi*, 6 (1), 24-31.
- Rahmayani, M. (2018). Analisis Clustering Tingkat Keparahan Penyakit Pasien Menggunakan Algoritma K-Means (Studi Kasus di Puskesmas Bandar Seikijang). *Jurnal Inovasi Teknik Informatika*, 1 (2), 40-44.
- Sari, Y., Primajaya, A., & Irawan, A. (2020). Implementasi Algoritma K-Means untuk Clustering Penyebaran Tuberkulosis di Kabupaten Karawang. *INOVTEK Polbeng-Seri Informatika*, 5 (2), 229-239.
- Toresa, D. (2020). Implementasi K-Means Terhadap Penyebaran Penyakit TBC di Riau Menggunakan Rapid Miner. *JUTIM (Jurnal Teknik Informatika Musirawas)*, 5 (1), 35-42.
- Wardhani, A., & Khrisna, A. (2016). Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien di Puskesmas Kajen Pekalongan. *J. Transformasi*, 14(1).
- Widiarina, W., & Wahono, R. S. (2015). Algoritma cluster dinamik untuk optimasi cluster pada algoritma k-means dalam pemetaan nasabah potensial. *Journal of Intelligent Systems*, 1(1), 33-36.
- Yuan, C., & Yang, H. (2019). Penelitian tentang metode pemilihan nilai K dari Algoritma Clustering K-Means. *J.2(2)*, 226-235.
- Zakiyyah, A. M., Rosyidah, U. A., & Margaretha, M. S. (2021). *Laporan Penelitian: Analisis Technology Readiness Sebagai Upaya Mitigasi Teknologi Pada UMKM Di Jember*. Jember: LPPM Universitas Muhammadiyah Jember.

# Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Penyakit Tbc (Studi Kasus: Puskesmas Di Kabupaten Jember)

## ORIGINALITY REPORT

10%

SIMILARITY INDEX

10%

INTERNET SOURCES

8%

PUBLICATIONS

1%

STUDENT PAPERS

## PRIMARY SOURCES

1	<a href="http://ejournal.polbeng.ac.id">ejournal.polbeng.ac.id</a> Internet Source	4%
2	<a href="http://docplayer.info">docplayer.info</a> Internet Source	2%
3	<a href="http://lifeeternityevangelicalministries.com">lifeeternityevangelicalministries.com</a> Internet Source	1%
4	Yanti Puspita Sari, Aji Primajaya, Agung Susilo Yuda Irawan. "Implementasi Algoritma K-Means untuk Clustering Penyebaran Tuberkulosis di Kabupaten Karawang", INOVTEK Polbeng - Seri Informatika, 2020 Publication	1%
5	<a href="http://tirto.id">tirto.id</a> Internet Source	1%
6	<a href="http://id.scribd.com">id.scribd.com</a> Internet Source	1%

---

Exclude quotes Off

Exclude matches < 20 words

Exclude bibliography On