



## Klasifikasi Keluhan Pasien terhadap Data Rekam Medis Pasien dengan Menggunakan Metode K Means

M. Agung Vafky Ideal<sup>1✉</sup>

<sup>1</sup>Independent Researcher

[mhdagung47@gmail.com](mailto:mhdagung47@gmail.com)

### Abstract

Patient complaints are the effects of the body being infected by a disease. Complaints experienced by patients can be caused by several factors, factors that cause patient complaints such as internal factors and external factors. Diseases caused by internal factors such as diseases caused by the influence of genes present in the patient. Diseases caused by external factors such as diseases caused by environmental conditions around the patient. Knowing the factors that cause patient complaints, the puskesmas can prevent a disease effectively and provide public health services on target. The data used in this study is data on patient complaints that are in the patient's medical record data. Medical record data is a file containing records of biodata and details of services that have been received by patients. This study aims to produce groups of patient complaint data and determine the parameters for each patient complaint. Classification of patient complaints can provide analysis and consideration of the determination of the puskesmas in providing health services and appropriate handling based on patient complaint data. In this study, the data used were patient complaints data within a period of 6 months consisting of 72 categories of patient complaints. The k-means method groups data with system partitions. The result is to classify data using the k-means method which produces data groups based on complaints so that they can be used as reference material for making more appropriate decisions in handling public health. This study produces a classification based on patient complaints so that the puskesmas can find patterns of information in the data and can help the puskesmas in controlling diseases that exist in the community.

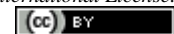
Keywords: K-Means, Medical Records, Data Mining, Classification, Complaints.

### Abstrak

Keluhan pasien merupakan efek dari tubuh yang sedang terjangkit oleh suatu penyakit. Keluhan yang dialami pasien dapat disebabkan oleh beberapa faktor. Faktor penyebab keluhan penyakit pasien terdiri atas faktor internal dan faktor eksternal. Penyakit yang disebabkan oleh faktor internal seperti penyakit yang disebabkan oleh pengaruh gen yang ada pada pasien. Penyakit yang disebabkan oleh faktor eksternal seperti penyakit yang disebabkan oleh kondisi lingkungan sekitar dari pasien. Mengetahui faktor penyebab keluhan pasien, maka pihak puskesmas dapat melakukan pencegahan suatu penyakit dengan efektif dan memberikan pelayanan kesehatan masyarakat dengan tepat sasaran. Data yang digunakan pada penelitian ini adalah data keluhan pasien yang berada pada data rekam medis pasien. Data rekam medis adalah berkas yang berisikan pencatatan dari biodata dan rincian pelayanan yang telah diterima pasien. Penelitian ini bertujuan untuk menghasilkan kelompok data keluhan pasien dan Mengetahui parameter pada setiap keluhan pasien. Klasifikasi keluhan pasien dapat memberikan Analisa dan pertimbangan terhadap penentuan pihak puskesmas dalam memberikan pelayanan kesehatan dan penanganan yang tepat berdasarkan data keluhan pasien. Pada penelitian ini data yang digunakan adalah data keluhan pasien dalam jangka waktu 6 bulan yang terdiri dari 72 kategori keluhan pasien. Data mining adalah suatu cara pengolahan data yang bertujuan untuk menemukan atau menggali informasi dari data yang diolah, salah satu metodenya yaitu metode k means. Metode k-means mengelompokkan data dengan system partisi. Hasilnya adalah melakukan klasifikasi data dengan metode k-means yang menghasilkan kelompok data berdasarkan keluhan sehingga dapat menjadi bahan acuan untuk mengambil keputusan yang lebih tepat dalam penanganan Kesehatan masyarakat. penelitian ini menghasilkan klasifikasi berdasarkan keluhan pasien sehingga pihak puskesmas dapat menemukan pola informasi yang ada pada data dan dapat membantu pihak puskesmas dalam pengendalian penyakit yang ada pada masyarakat.

Kata kunci: K-Means, Rekam Medis, Data Mining, Klasifikasi, Keluhan.

*JSISFOTEK is licensed under a Creative Commons 4.0 International License.*



### 1. Pendahuluan

Setiap harinya puskesmas padusunan menerima pasien yang cukup banyak. Penyakit yang banyak dan dengan gejala yang berbeda-beda dari masyarakat menyebabkan pihak puskesmas kesulitan dalam menentukan penyuluhan atau sosialisasi yang tepat dalam suatu penyakit yang ditimbulkan. Penyuluhan

yang tepat akan meminimalisir kemungkinan masyarakat terjangkit suatu penyakit. Keluhan yang diderita pasien dapat menjadi acuan dalam mengetahui penyebab dari suatu penyakit. Apakah penyakit tersebut disebabkan oleh faktor internal, atau faktor eksternal. Faktor internal seperti suatu penyakit yang disebabkan oleh suatu kondisi yang abnormal pada pasien itu sendiri. Faktor eksternal seperti penyakit

yang disebabkan oleh lingkungan tinggal pasien, ditularkan dari orang lain dan lain sebagainya.

Salah satu solusi yang dilakukan dalam mengatasi permasalahan di atas adalah dengan menggunakan Teknik data mining. Data mining yang digunakan yaitu clustering dengan menggunakan algoritma k-means clustering. Hasil dari penerapan algoritma clustering adalah berupa cara menentukan kelompok gejala penyakit yang terjadi pada masyarakat sehingga dari clustering tersebut pihak puskesmas dapat menentukan penyuluhan yang efektif dan efisien agar dapat mencegah penyebab dari suatu gejala penyakit.

Puskesmas merupakan salah satu lembaga pemerintah yang bergerak dibidang pelayanan kesehatan masyarakat pada tingkat kecamatan [1]. Salah satu peran puskesmas yaitu dalam menopang kinerja dari Lembaga Kesehatan di atasnya yaitu rumah sakit, sebagai upaya pencegahan dan penanggulangan masalah kesehatan masyarakat. Dalam upaya peningkatan kualitas pelayanan kesehatan yang lebih baik ditingkat puskesmas pada khususnya.

Rekam medis atau bisa disebut dengan ICD (Internasional Classification Diseases) adalah rekaman atau catatan dari pasien yang telah melakukan pengobatan di rumah sakit atau puskesmas. Bahasa medis yang biasa dilakukan oleh dokter dalam melakukan identifikasi suatu penyakit lalu memberikan tindakan atas penyakit tersebut berupa bahasa kedokteran (rekam medis) yang selanjutnya dikodekan menjadi kode-kode ICD. Kode ini adalah bahasa standar yang biasa digunakan oleh dokter pada umumnya walaupun bukan dokter spesialis, dalam membaca kode ini sesuai dengan aturan yang telah ditetapkan [2].

Knowledge Discovery in Database In Database (KDD) adalah suatu metode atau model yang digunakan untuk memperoleh pengetahuan yang tersembunyi dari database yang ada, pengetahuan yang didapatkan dapat menjadi acuan dalam mengambil suatu keputusan [3]. Database yang umumnya hanya dipakai untuk menyimpan suatu data namun dengan KDD kita dapat menemukan suatu pengetahuan tersembunyi didalamnya, dari pengetahuan tersebut dapat digunakan sebagai patokan atau pedoman dalam menganalisa sesuatu atau melakukan suatu tindakan. Secara garis besar KDD terdiri dari beberapa langkah yaitu data set cleaning, data integration, data selection, transformation of data, dan data mining.

Data mining dapat digunakan untuk menggali informasi dan nilai-nilai tersembunyi dari suatu kumpulan data, dimana informasi ini tidak dapat ditemukan apabila dilakukan secara manual [4]. Data mining merupakan salah satu langkah yang penting

pada proses KDD. Tahapan ini juga bisa disebut proses inti dari KDD. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari database yang besar. Salah satu metode dalam KDD adalah K-Means.

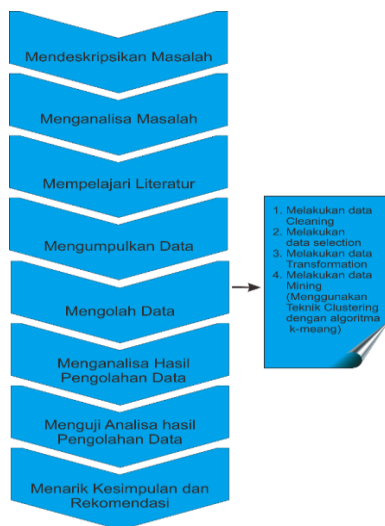
K-Means adalah metode partisi yang sering digunakan dalam penyesuaian pada machine study dan analisa suatu pola data, algoritma ini sangat terpengaruh pada centroid awal, tapi tidak dapat menjamin sampai menemukan solusi akhir hal ini dikarenakan centroid awal di cari secara acak untuk cluster yang diberikan [5]. K-means clustering sebagai salah satu metode data clustering non-hirarki yang mempartisi data yang tersedia ke dalam bentuk satu atau beberapa cluster, sehingga dapat menghasilkan kelompok data yang memiliki karakteristik atau cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. K-means clustering merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya.

Algoritma K-Means Clustering sebelumnya telah dilakukan dalam mengelompokkan data rekam medis berdasarkan jenis penyakit di Poli Klinik PT. Inecda dengan menggunakan metode k-means clustering. Data yang dipakai yaitu data rekam medis 600 pasien, Sehingga mendapatkan klasifikasi data berdasarkan wilayah, jenis penyakit, dan umur [2]. Segmentasi tumor otak dengan menggunakan metode k-means, penelitian ini dilakukan pada data MRI yang telah dianalisis, menghasilkan tingkat ke akurasian yang tinggi dalam mendiagnosa tumor [6]. Kinerja agen WFH (work from home) dan WFO (work from office) studi kasus PT. infomedia telkom dengan menggunakan metode k-means. Data yang digunakan yaitu 80% data agen WFH dan 20% data agen WFO, dihasilkan data yang menunjukkan perbedaan kinerja dari karyawan yang bekerja dari rumah dan di kantor [7]. Klasterisasi tingkat kehadiran dosen dengan menggunakan metode k-means clustering, data yang dipakai adalah data pertemuan yang dilakukan selama proses pembelajaran oleh dosen non PNS yang mengampu matakuliah. Penelitian ini menghasilkan pengelompokan dosen yaitu 72 matakuliah jarang melakukan pertemuan (4.7650%), 69 matakuliah sedang dalam melakukan pertemuan (4.5665%), dan 1370 matakuliah rajin melakukan pertemuan (90.6684%). Dosen non PNS rajin masuk dengan tingkat kehadiran 12-16 kali pertemuan per semester [8]. Akurasi pemetaan kelompok belajar siswa menuju prestasi dengan menggunakan metode k-means, data

yang diteliti yaitu nilai rapor siswa Kelas IX.C di SMP Pembangunan Laboratorium UNP. Hasil penelitian ini memperoleh 3 Cluster siswa yaitu siswa Sangat Berprestasi, Berprestasi dan Kurang Berprestasi. Penelitian ini bisa dijadikan sebagai pedoman bagi guru dalam pengambilan keputusan di Kelas IX.C [9].

## 2. Metodologi Penelitian

Pada penelitian ini menggunakan kerangka kerja yang dapat dilihat pada Gambar 1.



Gambar 1. Kerangka kerja penelitian

Berdasarkan kerangka kerja penelitian pada Gambar 1 menjelaskan beberapa kerangka kerja atau tahapan yang akan dilakukan dalam penelitian ini.

### 2.1. Mendeskripsikan masalah

Mendeskripsikan masalah adalah menjelaskan tentang masalah dalam penelitian secara terstruktur dan sistematis untuk mendukung dalam pengambilan keputusan yang lebih baik dan tepat. Masalah yang dideskripsikan dalam penelitian ini adalah klasifikasi keluhan pasien.

### 2.2. Menganalisa Masalah

Menganalisa masalah adalah tahap di mana peneliti menganalisa permasalahan yang ada sebagai langkah untuk dapat memahami masalah yang ada. Belum adanya klasifikasi keluhan pasien selama ini. Penerapan data mining teknik clustering menggunakan algoritma K-means clustering membantu dalam menentukan kelompok dari keluhan pasien yang ada.

### 2.3. Mempelajari Literatur

Tahap ini akan dicari dan dikumpulkan serta dipelajari sejumlah literatur mengenai teori dan konsep yang mendukung penyelesaian permasalahan dalam penelitian. Literatur yang digunakan berupa buku

referensi atau buku penunjang, jurnal internasional dan nasional serta konsep-konsep yang mendukung dalam menyelesaikan penelitian ini.

### 2.4. Mengumpulkan Data

Pengumpulan data dilakukan observasi yaitu pengamatan secara langsung di tempat penelitian sehingga permasalahan yang ada dapat diketahui secara jelas. Kemudian dilakukan interview atau wawancara dengan staf atau tenaga kesehatan yang ada di Puskesmas Padusunan.

### 2.5. Mengolah Data

Data yang diperlukan dalam penelitian selesai dikumpulkan, langkah selanjutnya adalah melakukan pengolahan data dengan beberapa tahap diantaranya adalah :

- Data cleaning, dilakukan untuk menghilangkan data yang tidak konsisten atau tidak relevan.
- Data selection, dilakukan untuk mengambil data yang sesuai untuk dianalisa. Pada penelitian ini atribut yang digunakan berupa data keluhan pasien perperiode (perbulan).
- Data transformation, agar data dapat diolah menggunakan algoritma k-means
- Data mining, setelah data ditransformasikan maka data tersebut dapat diproses menggunakan algoritma k-means clustering.

Tahapan disajikan dalam *pseudocode* pada Algoritma 1.

#### Algoritma 1. K-Means Clustering

```

    Input :x,k,n,s
    Output: C
    Procedure
    C←0;
    For i=1 to n
    Si←RandomSample(x,s)
    (ai,ci)←kmeans(Si,k)
    C←AppendRows(C,ci)
    End for
    Return Cfinal←kmeans(C,k)
    End procedure
  
```

### 2.6. Menganalisa hasil pengolahan data

Tahap ini dilakukan analisa terhadap hasil pengolahan data berupa cluster calon penerima beasiswa kurang mampu yang telah diolah dengan menggunakan algoritma k-means clustering.

### 2.7. Menguji analisa hasil pengolahan data

Hasil analisa pengolahan data yang dilakukan dengan teknik clustering menggunakan algoritma k-means clustering selanjutnya dilakukan pengujian. Pengujian dilakukan terhadap data yang diolah secara manual berdasarkan langkah-langkah algoritma k-means clustering, kemudian dibandingkan dengan hasil perhitungan menggunakan WEKA.

2.8. Menarik kesimpulan dan rekomendasi

Langkah terakhir setelah semua proses selesai, maka dapat ditarik suatu kesimpulan terhadap masalah yang diteliti. Selanjutnya diberikan rekomendasi dengan tujuan kelemahan-kelemahan dapat dihilangkan dan solusi kedepannya dapat diimplementasikan sehingga tujuan yang diharapkan dapat dicapai.

3. Hasil dan Pembahasan

3.1 Input Data Keluhan Pasien

Pada penelitian ini data yang digunakan adalah data keluhan pasien yang berjumlah 73 kategori dalam jangka waktu 6 bulan dari bulan Februari – Juli. Data tersebut dapat dilihat pada Tabel 1.

Tabel 1. Keluha Pasien Perbulan

Keluhan	Februari	Maret	April	Mei	Juni	Juli
Amandel	6	8	4	2	3	3
Asam Lambung	5	7	8	6	10	9
Asam Urat	4	6	4	2	3	5
BAB Berdarah	2	9	6	5	1	4
BAB Tidak Lancar	2	7	6	6	1	2
Badan Dingin	6	9	6	7	14	7
Badan Pegal	1	7	9	6	7	10
BAK Tidak Lancar	6	11	2	4	5	2
Batuk	6	16	5	6	17	9
Batuk Berdahak	7	5	8	7	7	7

Pada Tabel 1. Dapat kita lihat, kolom pertama berisikan kategori keluhan pasien. Data direkap selama 6 bulan yang dimulai dari bulan february sampai bulan juli

3.2 Tentukan Jumlah Cluster

Tahapan berikutnya yang akan dilakukan sebelum pengolahan data adalah menentukan jumlah cluster. Jumlah cluster adalah jumlah kelompok data yang akan dihasilkan pada penelitian ini. Pada penelitian ini akan mengelompokkan data keluhan pasien menjadi 3 cluster.

3.3 Proses Iterasi Pertama

Proses iterasi pertama adalah tahapan yang dilakukan untuk mendapatkan kelompok data keluhan pasien berdasarkan jarak terdekat data dengan pusat klaster yang telah ditentukan. Pada tahapan iterasi pertama terdiri dari beberapa tahapan yaitu menentukan centroid awal, menghitung jarak, dan mengelompokkan data.

a. Tentukan pusat cluster (centroid) awal

Pada tahapan ini, dalam menentukan centroid awal diambil secara acak dari data keluhan pasien yang akan diolah. Data yang akan dijadikan centroid awal dapat dilihat pada tabel 4.4. Data centroid awal terdiri dari C1, C2, dan C3, berikut adalah Tabel 2. centroid awal.:

Tabel 2. Centroid Awal

Pusat	Nama	Februari	Maret	April	Mei	Juni	Juli
C1	Mual	18	5	15	7	8	19
C2	Sakit persendian	5	8	14	3	3	2
C3	Kaki gatal	7	7	3	3	8	6

Pada Tabel 2. berisikan centroid awal yang ditentukan. Centroid awal ini diambil secara acak dari kategori keluhan pasien. Centroid awal yang dipilih yaitu mual, sakit persendian, dan kaki gatal.

b. Menghitung jarak iterasi pertama

Proses iterasi pertama akan menghasilkan kelompok data pada pengolahan data pertama. Menggunakan rumus euclidean untuk menghitung jarak data terhadap centroid awal. Berikut adalah tahapan perhitungan iterasi pertama :

$$D(a, b) = \sqrt{(xa - yb)^2 + (xa - yb)^2 + (xn - yn)^2} \quad (1)$$

Perhitungan jarak keluhan pertama dengan pusat cluster pertama. Perhitungan ini dilakukan juga pada data keluhan pasien ke 3 sampai data ke 72. Hasil

perhitungan jarak keluhan pasien dengan ketiga pusat cluster awal. Berikut adalah hasil iterasi 1 pada Tabel 3.

Tabel 3. Jarak Setiap Data Keluhan pada Iterasi 1

Ke	Cluster		
	C1	C2	C3
1	24,08318916	10,14889157	6,16441400
2	18,08314132	12,00000000	7,14142842
3	23,74868417	10,72380529	6,16441400
4	25,11971337	9,27361849	9,74679434
5	26,07680962	9,32737905	10,39230485
6	20,51828453	15,09966887	8,18535277
7	20,29778313	11,44552314	9,89949493
8	25,61249695	12,60952021	6,63324958
9	23,38803113	20,00000000	13,60147051
10	17,74823935	10,29563014	6,85565460

Pada Tabel 3. Berisikan hasil dari perhitungan jarak dari masing-masing kategori keluhan pasien terhadap *centroid* awal. Pada kolom c1 merupakan jarak kategori terhadap *centroid* awal c1. Kolom c2 berisikan hasil perhitungan jarak data keluhan terhadap *centroid* awal c2. Kolom c3 berisikan jarak data keluhan terhadap *centroid* awal c3.

c. Hasil Pengelompokan Iterasi Pertama

*Clustering* objek dengan memasukkan setiap objek ke dalam *cluster* berdasarkan pada jarak minimumnya. Suatu data akan menjadi anggota dari suatu *cluster* (C1, C2 atau C3) yang memiliki jarak atau nilai terkecil dari pusat *cluster*-nya, misalnya untuk data pertama, jarak terkecil ada pada *cluster* yang ke-3 yaitu 6,164414003, maka keluhan pertama akan menjadi anggota pada *cluster* 2. Berikut adalah Tabel 4. hasil iterasi 1.

Tabel 4. Hasil Pengelompokan Iterasi 1

Keluhan ke	Cluster			Anggota Cluster
	C1	C2	C3	
1	24,08318916	10,14889157	6,164414003	3
2	18,08314132	12	7,141428429	3
3	23,74868417	10,72380529	6,164414003	3
4	25,11971337	9,273618495	9,746794345	2
5	26,07680962	9,327379053	10,39230485	2
6	20,51828453	15,09966887	8,185352772	3
7	20,29778313	11,44552314	9,899494937	3
8	25,61249695	12,60952021	6,633249581	3
9	23,38803113	20	13,60147051	3
10	17,74823935	10,29563014	6,8556546	3

Pada Tabel 4. merupakan hasil pengelompokan *cluster* dari data keluhan. Pengelompokan berdasarkan jarak terdekat dari hasil perhitungan jarak terdekat dari ketiga *centroid* awal.

3.4 Proses Iterasi Kedua

Proses iterasi kedua adalah tahapan yang dilakukan untuk mendapatkan kelompok data keluhan pasien berdasarkan jarak terdekat data dengan pusat klaster yang baru. Pada tahapan iterasi kedua terdiri dari beberapa tahapan yaitu menentukan *centroid* kedua, menghitung jarak, dan mengelompokkan data.

Pusat C1 yang baru adalah (15,125; 4,875; 12,375; 12,75; 11,375; 16,5). Perhitungan ini dilakukan juga pada penentuan pusat C2 dan C3 yang baru. Berikut adalah hasil penentuan pusat baru pada Tabel 5.

Tabel 5. Hasil Penentuan Pusat Centroid Baru

Pusat	Februari	Maret	April	Mei	Juni	Juli
C1	15,125	4,875	12,375	12,750	11,375	16,500
C2	4,800	5,100	10,200	4,300	4,200	3,600
C3	4,407	6,962	4,166	5,129	5,481	6,240

Pada Tabel 5. Berisikan hasil perhitungan pencarian *centroid* yang baru. Pada bari C1 merupakan *centroid*

awal c1. Pada bari C2 merupakan *centroid* awal c2. Pada bari C3 merupakan *centroid* awal c3.

a. Menghitung Jarak Iterasi Kedua

Melakukan penentuan pusat *centroid* yang baru, maka proses selanjutnya adalah melakukan iterasi 2. Iterasi dua langkahnya sama dengan iterasi 1. Iterasi dua tentunya dihitung dengan menggunakan hasil dari penentuan *centroid* yang baru. Hasil dari iterasi 2 dapat dilihat pada Tabel 6.

Tabel 6. Hasil Iterasi 2

Keluhan ke	Cluster		
	C1	C2	C3
1	23,04614935	7,441773982	5,485826869
2	15,16162920	8,612781200	6,620803307
3	22,65226258	6,969935437	4,314257344
4	23,53454907	7,181921748	6,199002858
5	24,10134851	6,691786010	6,926988988
6	16,46587380	12,213926480	9,310580886
7	17,95480437	8,436824047	7,222767921
8	23,57700151	10,333440860	6,558982881
9	19,87146195	18,525118080	15,031508120
10	15,08724627	6,031583540	5,625816957

Pada Tabel 6 Berisikan hasil dari perhitungan jarak dari masing-masing kategori keluhan pasien terhadap *centroid* awal. Pada kolom c1 merupakan jarak kategori terhadap *centroid* awal c1. Kolom c2 berisikan hasil perhitungan jarak data keluhan terhadap *centroid* awal c2. Kolom c3 berisikan jarak data keluhan terhadap *centroid* awal c3.

b. Hasil Pengelompokan Iterasi Kedua

Pada tahapan ini caranya sama dengan pengelompokan iterasi pertama. *Clustering* objek dengan memasukkan setiap objek ke dalam *cluster* berdasarkan pada jarak minimumnya. Suatu data akan menjadi anggota dari suatu *cluster* (C1, C2 atau C3) yang memiliki jarak atau nilai terkecil dari pusat *cluster*-nya. Berikut adalah Tabel 7. hasil pengelompokan iterasi kedua.

Tabel 7. Hasil Pengelompokan Iterasi 2

Ke	Cluster			Anggota Cluster
	C1	C2	C3	
1	23,04614935	7,441773982	5,485826869	3
2	15,16162920	8,612781200	6,620803307	3
3	22,65226258	6,969935437	4,314257344	3
4	23,53454907	7,181921748	6,199002858	3
5	24,10134851	6,691786010	6,926988988	2
6	16,46587380	12,213926480	9,310580886	3
7	17,95480437	8,436824047	7,222767921	3
8	23,57700151	10,333440860	6,558982881	3
9	19,87146195	18,525118080	15,031508120	3
10	15,08724627	6,031583540	5,625816957	3

Pada Tabel 7. merupakan hasil pengelompokan *cluster* dari data keluhan. Pengelompokan berdasarkan jarak terdekat dari hasil perhitungan jarak terdekat dari ketiga *centroid* awal.

### 3.5 Hasil K-Means Clustering

Proses iterasi K-means berhenti ketika hasil iterasi tidak berubah dengan hasil iterasi sebelumnya. Proses

iterasi akan terus dilakukan sampai iterasi tidak berubah dari nilai iterasi sebelumnya. Berikut adalah hasil dari proses algoritma K-means yang telah selesai pada Tabel 8.

Tabel 8. Hasil Clustering Data Keluhan Pasien

Keluhan	Februari	Maret	April	Mei	Juni	Juli	Hasil
Amandel	6	8	4	2	3	3	C2
Asam lambung	5	7	8	6	10	9	C1
Asam urat	4	6	4	2	3	5	C3
BAB berdarah	2	9	6	5	1	4	C3
BAB tidak lancar	2	7	6	6	1	2	C3
Badan dingin	6	9	6	7	14	7	C3
Badan Pegal	1	7	9	6	7	10	C3
BAK tidak lancar	6	11	2	4	5	2	C2
Batuk	6	16	5	6	17	9	C2
Batuk berdahak	7	5	8	7	7	7	C3

Berdasarkan hasil cluster dari Tabel 8 bahwa data keluhan pasien kita dapat menyimpulkan bahwa, C1 merupakan keluhan penyakit yang sering terjadi dengan intensitas tinggi sehingga pihak puskesmas memberikan perhatian khusus pada keluhan tersebut yaitu dengan cara memberikan penyuluhan luar ruangan, penyuluhan dalam ruangan. C2 merupakan keluhan yang terjadi dengan intensitas sedang sehingga tindakan yang dapat dilakukan puskesmas yaitu dengan cara penyuluhan dalam ruangan, C3 merupakan keluhan yang terjadi dengan intensitas rendah sehingga puskesmas hanya melakukan arahan atau peringatan pada saat pasien melakukan pengobatan.

### 4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan hasil *cluster* dapat menjadi patokan atau titik acuan bagi pihak puskesmas dalam memberikan pelayanan kesehatan pada masyarakat dan pencegahan penyakit yang lebih efektif.

### Daftar Rujukan

[1] Safitri, N. E. (2019). Implementasi Peraturan Menteri Kesehatan Republik Indonesia Nomor 75 Tahun 2014 Tentang Pusat Kesehatan Masyarakat Di Puskesmas Sukamakmur Labuhan Batu (Doctoral dissertation). <http://repository.umsu.ac.id/handle/123456789/3326>

[2] Ordila, R., Wahyuni, R., Irawan, Y., & Sari, M. Y. (2020). Penerapan Data Mining Untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit Dengan Algoritma Clustering (Studi Kasus: Poli Klinik Pt. Inecda). *Jurnal Ilmu Komputer*, 9(2), 148-153. DOI : <https://doi.org/10.33060/JIK/2020/Vol9.Iss2.181>

[3] Ghazal, M. M., & Hammad, A. (2022). Application of knowledge discovery in database (KDD) techniques in cost overrun of construction projects. *International Journal of Construction Management*, 22(9), 1632-1646. <https://doi.org/10.1080/15623599.2020.1738205>

[4] Manochandar, S., Punniyamoorthy, M., & Jeyachitra, R. K. (2020). Development of new seed with modified validity measures for k-means clustering. *Computers & Industrial Engineering*, 141, 106290. DOI : <https://doi.org/10.1016/j.cie.2020.106290>

[5] Khan, A. R., Khan, S., Harouni, M., Abbasi, R., Iqbal, S., & Mehmood, Z. (2021). Brain tumor segmentation using K-means clustering and deep learning with synthetic data augmentation for classification. *Microscopy Research and Technique*, 84(7), 1389-1399. DOI : <https://doi.org/10.1002/jemt.23694>

[6] Jaja, J., Priatna, N., & Ardan, T. S. (2021). Implementation of Data Mining Technique for Performance of WFH and WFO Agents Using the K-Means Method Case Study Study of PT. Infomedia Telkom Consumer Profiling Services. *Budapest International Research in Exact Sciences (BirEx) Journal*, 3(2), 117-125. DOI : <https://doi.org/10.33258/birex.v3i2.1810>

[7] Virgo, I., Defit, S., & Yuhandri, Y. (2020). Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering. *Jurnal Sistim Informasi dan Teknologi*, 23-28. DOI : <https://doi.org/10.37034/jsisfotek.v2i1.17>

[8] Dewi, S., Defit, S., & Yuhandri, Y. (2021). Akurasi Pemetaan Kelompok Belajar Siswa Menuju Prestasi Menggunakan Metode K-Means. *Jurnal Sistim Informasi dan Teknologi*, 28-33. DOI : <https://doi.org/10.37034/jsisfotek.v3i1.40>

[9] Yuhafizar, Santosa B., Eddy, I. K. P., & Suprpto, Y. K. (2013). Combination of Cluster Method for Segmentation of Web Visitors. *TELKOMNIKA*, 11(1), 207-214. <http://dx.doi.org/10.12928/telkomnika.v11i1.906>