

Klasterisasi Data Kesehatan Penduduk untuk Menentukan Rentang Derajat Kesehatan Daerah dengan Metode K-Means

Nielza Atthina

Jurusan Teknik Informatika
Universitas Islam Indonesia
Yogyakarta
09523458@students.uui.ac.id

Lizda Iswari

Jurusan Teknik Informatika
Universitas Islam Indonesia
Yogyakarta
Lizda.iswari@uui.ac.id

Abstract—Salah satu indikator untuk menilai tingkat keberhasilan pembangunan negara adalah tingkat capaian negara dalam memberikan jaminan bidang kesehatan untuk masyarakatnya. Pemerintah Indonesia, melalui Departemen Kesehatan, menetapkan sejumlah indikator sebagai tolak ukur kemajuan pembangunan kesehatan. Setiap tahun Departemen Kesehatan mengumpulkan data kesehatan penduduk di berbagai level unit daerah untuk diolah sehingga dihasilkan ranking Provinsi dan Kabupaten/Kota Sehat. Metode pengolahan terhadap data kesehatan yang sudah terkumpul tersebut masih berbasis pada teknik statistik dasar, dimana memiliki kelemahan dalam hal konsistensi data dan minimnya informasi tentang hubungan antar data. Data mining dilihat dari sisi teknik pengolahan data menyediakan sejumlah algoritma yang dapat digunakan untuk menggali informasi tersembunyi dari kumpulan data yang multidimensi. Penelitian ini membahas tentang pemanfaatan algoritma k-Means untuk mengklaster atau mengelompokkan kabupaten-kabupaten di Provinsi Jawa Tengah dan DIY berdasarkan kemiripan nilai lima buah ukuran indikator mortalitas derajat kesehatan daerah, yaitu angka kelahiran kasar, angka kematian kasar, angka kematian bayi, angka kematian balita, dan angka kematian ibu. Secara umum, output penelitian berupa hasil klasterisasi dapat digunakan sebagai sebuah acuan untuk menggambarkan distribusi pengelompokan kabupaten berdasarkan kondisi profil kesehatan daerah tersebut.

Keywords—derajat kesehatan, k-Means, klastering, mortalitas

I. PENDAHULUAN

Salah satu indikator untuk menilai tingkat keberhasilan pembangunan negara dapat dilihat dari tingkat capaian negara tersebut dalam memberikan jaminan bidang kesehatan. Di Indonesia, pemerintah melalui Departemen Kesehatan menetapkan sejumlah indikator sebagai tolak ukur kemajuan pembangunan kesehatan di berbagai level unit daerah di Indonesia, dari tingkat provinsi, kabupaten, hingga kecamatan.

Setiap tahun Departemen Kesehatan, melalui Dinas Kesehatan yang tersebar di daerah, mengumpulkan data

kesehatan penduduk untuk diolah sehingga dihasilkan ranking Provinsi dan Kabupaten/Kota Sehat. Hasil olahan dan perankingan ini merupakan acuan penting, seperti bagi Pemerintah Daerah (Pemda) untuk membuat program intervensi yang lebih tepat sasaran, sebagai bahan advokasi agar Pemda lebih terpacu untuk menaikkan ranking kesehatan daerahnya, sebagai landasan pemerintah pusat untuk merumuskan Daerah Bermasalah Kesehatan Berat/Khusus (DBKKB), sebagai dasar untuk menentukan alokasi dana bantuan kesehatan dari pemerintah pusat ke daerah, dan juga berperan membantu Kementerian Negara Pembangunan Daerah Tertinggal (KMPDT) untuk membangun kabupaten/kota [1].

Metode pengolahan data indikator kesehatan yang selama ini diterapkan masih berbasis pada teknik statistik dasar, yaitu perhitungan yang didasarkan pada hasil rata-rata seluruh indikator atau didasarkan pada distribusi data [2]. Ternyata output yang dihasilkan dari metode ini masih memiliki sejumlah permasalahan terutama dalam hal konsistensi data. Dalam hal ini, daerah yang memiliki ranking lebih tinggi belum tentu memiliki semua nilai indikator yang lebih baik dibandingkan dengan daerah di ranking yang lebih rendah. Permasalahan lain terletak pada minimnya informasi tentang hubungan antar indikator. Seperti pada derajat kesehatan daerah belum ada analisis yang dilakukan untuk mengukur, misal, pengaruh status gizi masyarakat terhadap kejadian mortalitas (kematian) maupun morbiditas (kesakitan).

Data mining dilihat dari sisi teknik pengolahan data menyediakan sejumlah algoritma yang dapat digunakan untuk menggali informasi tersembunyi dari kumpulan data yang multidimensi [3]. Selain itu, teknik data mining juga dapat digunakan untuk mengekstrak pola keterhubungan antar data. Salah satu algoritma data mining yang cukup populer digunakan baik dalam dunia bisnis, akademik, ataupun industri adalah algoritma k-Means. Algoritma ini bekerja dengan cara membagi data dalam sejumlah klaster untuk dianalisis faktor kesamaan (*similarity*) maupun ketidaksamaan (*dissimilarity*) yang melekat pada kumpulan data tersebut, dan kemudian dianalisis pola keterhubungan antar data [4].

Penelitian ini membahas tentang pemanfaatan algoritma k-Means untuk mengklaster atau mengelompokkan kabupaten-kabupaten di Provinsi Jawa Tengah dan DIY

berdasarkan kemiripan karakteristik daerah yang ditinjau dari nilai lima buah ukuran indikator mortalitas derajat kesehatan daerah, yaitu angka kelahiran kasar, angka kematian kasar, angka kematian bayi, angka kematian balita, dan angka kematian ibu. Secara umum, output penelitian berupa hasil klasterisasi dapat digunakan sebagai sebuah acuan untuk menggambarkan distribusi pengelompokan kabupaten berdasarkan kondisi profil kesehatan daerah tersebut.

II. SITUASI DERAJAT KESEHATAN

Derajat kesehatan masyarakat adalah gambaran tentang kondisi kesehatan yang terjadi secara umum di masyarakat. Untuk mengukur atau mengkuantifikasi kondisi kesehatan yang dihadapi suatu daerah, digunakanlah sejumlah indikator kesehatan. Di Indonesia, derajat kesehatan masyarakat diukur berdasarkan kondisi mortalitas (kematian), status gizi, dan morbiditas (kesakitan) [2]. Angka mortalitas terdiri atas: angka kematian neonatal, angka kematian bayi, angka kematian balita, dan angka harapan hidup, sedangkan morbiditas mengacu pada angka kesakitan sejumlah penyakit pada balita dan dewasa.

Dalam penelitian ini digunakan lima buah indikator mortalitas yang mempengaruhi tingkat kesehatan suatu daerah beserta dengan cara pengukurannya [5]:

I. Angka Kelahiran Kasar (CBR)

Angka Kelahiran Kasar (*Crude Birth Rate/CBR*) adalah angka yang menunjukkan banyaknya kelahiran pada tahun tertentu per 1000 penduduk pada pertengahan tahun yang sama.

Rumus :

$$CBR = \left(\frac{B}{P}\right) \times k \quad (1)$$

II. Angka Kematian Kasar (CDR)

Angka Kematian Kasar (*Crude Death Rate/CDR*) adalah jumlah kematian yang dicatat dalam 1 tahun per 1000 penduduk pada pertengahan tahun yang sama.

Rumus :

$$CDR = \left(\frac{D}{P}\right) \times k \quad (2)$$

III. Angka Kematian Bayi (IMR)

Angka Kematian Bayi (*Infant Mortality Rate/IMR*) adalah jumlah kematian bayi yang berumur kurang dari 1 tahun yang dicatat selama 1 tahun per 1000 kelahiran hidup pada tahun yang sama.

Rumus :

$$IMR = \frac{d_0}{B} \times k \quad (3)$$

IV. Angka Kematian Balita (FMR)

Angka kematian balita (*Under Five MortalityRate/FMR*) adalah jumlah kematian balita yang dicatat selama 1 tahun per 1000 kelahiran hidup pada tahun yang sama.

Rumus :

$$FMR = \left(\frac{b}{Bl}\right) \times k \quad (4)$$

V. Angka Kematian Ibu (MMR)

Angka kematian ibu (*Maternal Mortality Rate/MMR*) adalah jumlah kematian ibu sebagai akibat komplikasi kehamilan, persalinan, dan masa nifas yang dicatat selama 1 tahun per 1000 kelahiran hidup pada tahun yang sama.

Rumus :

$$MMR = \left(\frac{I}{B}\right) \times k \quad (5)$$

Dengan :

B = Jumlah lahir hidup yang dicatat selama 1 tahun

D = Jumlah kematian dicatat selama 1 tahun.

d_0 = Jumlah kematian bayi yang berumur kurang dari 1 tahun

b = Jumlah kematian balita yang dicatat selama 1 tahun

Bl = Jumlah penduduk balita pada tahun yang sama

P = Jumlah penduduk pada pertengahan tahun yang sama

K = Konstanta (1000)

I = Jumlah kematian ibu sebagai akibat komplikasi kehamilan, persalinan dan masa nifas yang dicatat selama 1 tahun

III. KLASTERISASI INDIKATOR KESEHATAN

Pencarian pengetahuan dalam data, atau dikenal juga sebagai *Knowledge discovery in Database (KDD)*, didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses *knowledge discovery* melibatkan hasil dari proses *data mining* (proses mengekstrak kecenderungan pola suatu data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami [6].

Pada dasarnya ada tiga tahapan penting dalam teknik pencarian pengetahuan dalam data, yaitu: (1) Pra-prosesing yang meliputi pengumpulan dan pengambilan data dari basis data (*data collection*), pembersihan data (*data cleaning*), dan pemilihan dan transformasi data (*data selection and transformation*), (2) Data mining, yaitu penggalian data menggunakan sejumlah metode, dan (3) Post-prosesing, yang meliputi evaluasi hasil dan memvisualisasikannya dalam bentuk yang mudah dipahami pengguna [7].

Sebagai contoh terdapat data kondisi kesehatan penduduk di 10 kabupaten di Jawa Tengah seperti tampak pada Tabel 1. Langkah pertama adalah pra-prosesing data, dalam hal ini data input dikonversi menjadi nilai indikator kesehatan daerah, yaitu CDR, CBR, IMR, FMR, dan MMR, sesuai rumus (1) – (5). Hasil perhitungan data indikator kesehatan dapat dilihat pada Tabel 2. Setelah diperoleh nilai indikator kesehatan, langkah berikutnya adalah transformasi data. Semua data indikator kesehatan akan dinormalisasi ke dalam rentang 0 – 1. Normalisasi data sangat dibutuhkan sebelum proses data mining, agar tidak ada parameter yang mendominasi dalam perhitungan jarak antar data [6]. Adapun rumus untuk normalisasi dapat dilihat pada rumus (6) dan hasil normalisasi data terlihat pada Tabel 3.

$$\text{Nilai Baru} = \left(\frac{\text{nilai asal} - \text{nilai min}}{\text{nilai max} - \text{nilai min}} \right) \quad (6)$$

Langkah kedua adalah data mining. Metode yang diterapkan adalah klustering dengan algoritma k-Means. Algoritma ini disusun atas dasar ide yang sederhana. Awalnya ditentukan berapa kluster yang akan dibentuk, misal 3 buah kluster. Proses klustering dimulai dengan memilih secara random obyek yang dijadikan sebagai titik pusat kluster, misal obyek yang terpilih adalah obyek ke-1, 2, 9 dan diberi nama C1, C2, dan C3, seperti tampak pada Tabel 4. Kemudian dihitung jarak setiap data dengan semua titik pusat kluster menggunakan perhitungan Euclidean Distance [6] seperti pada rumus (7).

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2} \quad (7)$$

Dengan:

d_{ik} = jarak antara data ke-i dengan titik pusat kluster ke-k

m = jumlah atribut

x_j = data ke-i

c_k = data pusat kluster ke-k

Suatu data akan menjadi anggota sebuah kluster, jika memiliki jarak yang paling pendek. Setiap data akan berada tepat dalam satu buah kluster. Hasil perhitungan jarak beserta penentuan kluster dapat dilihat pada Tabel 5. Selanjutnya, pusat kluster akan diperbaharui berdasarkan nilai rata-rata anggotanya. Adapun rumus untuk menghitung titik pusat kluster baru dapat dilihat pada rumus (8) dan hasil kluster baru dapat dilihat pada Tabel 6.

$$c_{kj} = \frac{\sum_{h=1}^p y_{hj}}{p}; y_{hj} = x_{ij} \in \text{cluster ke-k} \quad (8)$$

Dengan:

c_{kj} = Pusat Cluster

p = Jumlah semua anggota cluster

h = Jumlah awal anggota cluster (nilai =1)

y_{hj} = Jumlah data

Kemudian dilakukan lagi perhitungan jarak setiap data dengan semua pusat kluster baru seperti yang ditampilkan dalam Tabel 7. Proses ini terus berlangsung sampai nilai pusat kluster tidak berubah. Dalam contoh menggunakan 10 data kabupaten ini diperlukan dua kali iterasi untuk menghasilkan nilai pusat kluster yang tidak berubah seperti tampak pada Tabel 8.

Langkah ketiga adalah analisis dan visualisasi hasil. Dalam penelitian ini digunakan peta sebagai alat bantu untuk menggambarkan distribusi pengelompokan kabupaten berdasarkan kemiripan nilai indikator kesehatan daerah.

IV. ANALISIS DAN PERANCANGAN SISTEM

A. Analisis Kebutuhan Sistem

Analisis kebutuhan sistem dilakukan untuk mengetahui kebutuhan data input, kebutuhan proses, dan keluaran sistem. Namun sebelumnya, telah ditetapkan bahwa sistem yang dibangun ditujukan untuk 4 jenis pengguna (user), yaitu:

- Admin: merupakan user dengan hak akses tertinggi untuk mengelola data sistem, seperti menentukan hak akses pengguna dan aktivasi data.
- Staf: adalah user yang berperan dalam memasukkan data transaksi rutin indikator kesehatan tiap daerah.
- Pakar: adalah user yang berperan untuk menjalankan fungsi klusterisasi data dan memberikan analisis atau menyimpulkan hasil klustering.
- Pengunjung: adalah user umum yang memiliki akses untuk melihat hasil klusterisasi data indikator kesehatan daerah.

Adapun kebutuhan data masukan sistem terdiri atas:

- Data indikator kesehatan: jumlah penduduk, jumlah penduduk balita, jumlah kelahiran, jumlah kematian kasar, jumlah kematian bayi, jumlah kematian balita, dan jumlah kematian ibu.
- Data spasial: kabupaten-kabupaten Provinsi Jawa Tengah (Jateng) dan Daerah Istimewa Yogyakarta (DIY).
- Data pelengkap: tahun, username, dan password.

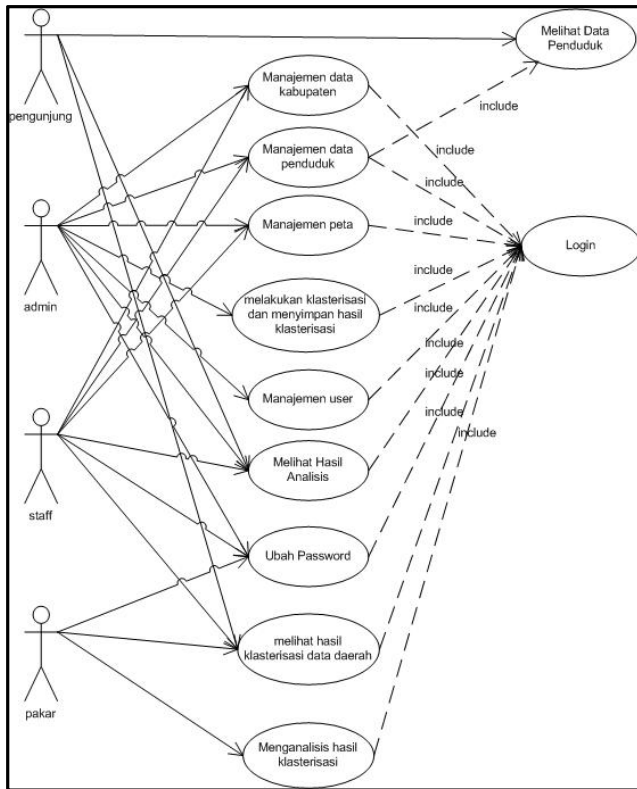
Keluaran sistem yang diharapkan adalah peta kabupaten di Provinsi Jateng dan DIY yang telah terbagi dalam 3 buah kluster beserta informasi derajat kesehatan daerah (CBR, CDR, IMR, FMR, dan MMR), dan informasi hasil analisis pakar. Untuk mencapai keluaran sistem tersebut di atas dibutuhkan sejumlah proses, berupa:

- Manajemen Data Daerah, yaitu fungsi untuk mengelola data kabupaten dan indikator kesehatan tiap kabupaten.
- Manajemen Pengguna, yaitu fungsi untuk mengelola akun beserta hak akses pengguna.
- Klusterisasi Data, yaitu fungsi untuk mengelompokkan/mengkluster data menggunakan algoritma k-Means.
- Analisis Hasil Kluster, yaitu fungsi untuk memasukkan analisis atau kesimpulan hasil klusterisasi data.
- Manajemen Peta, yaitu fungsi untuk menghubungkan data spasial kabupaten dengan hasil klusterisasi data.

Gambaran interaksi antara pengguna dengan sistem divisualisasikan dalam bentuk Use Case Diagram seperti pada Gambar 1, yaitu hak akses untuk setiap jenis pengguna.

Untuk user Admin memiliki keleluasan akses untuk semua fungsi, termasuk di dalamnya adalah menjalankan proses klusterisasi data. Jenis user Staf memiliki hak akses pada fungsi manajemen data daerah, dalam hal ini adalah mengelola data transaksi rutin yaitu data mentah yang akan dijadikan indikator kesehatan tiap kabupaten. Untuk user Pakar memiliki hak akses

untuk memasukkan analisis hasil klusterisasi. Sedangkan output dari proses klusterisasi berupa peta hasil kategorisasi data dapat dilihat oleh siapa saja, termasuk user umum berupa pengunjung web.



Gambar 1. Rancangan Pengguna dan Fungsi Sistem

B. Perancangan Basis Data

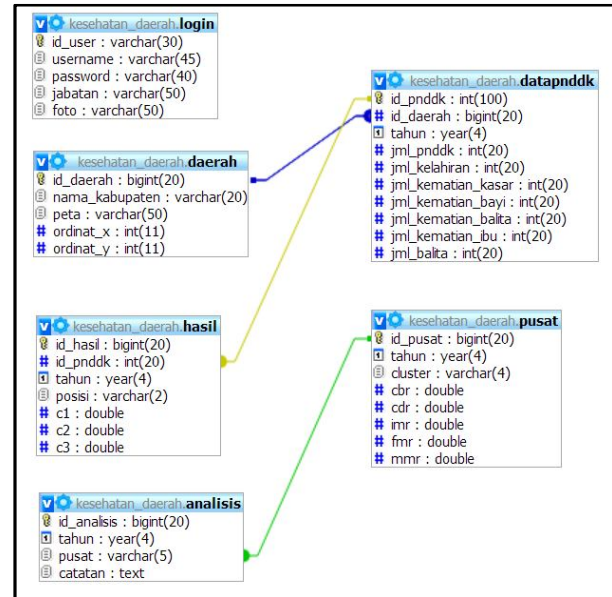
Untuk mengembangkan sistem klusterisasi indikator kesehatan daerah dibutuhkan sejumlah tabel yang berfungsi untuk menyimpan data relevan. Terdapat enam buah tabel yang dibutuhkan. Berikut adalah penjelasan singkat fungsi setiap tabel:

- Tabel Login: berfungsi untuk menyimpan data akun, password, dan jenis user.
- Tabel Daerah: berfungsi untuk menyimpan data spasial daerah beserta deskripsinya, seperti nama kabupaten dan posisi titik pusat (ordinat x dan ordinat y) daerah untuk ditampilkan di peta.
- Tabel Data Penduduk: berfungsi untuk menyimpan data mentah indikator kesehatan tiap daerah. Tabel ini memiliki relasi terhadap tabel Daerah.
- Tabel Hasil: berfungsi untuk menyimpan hasil klusterisasi indikator kesehatan daerah pada tahun tertentu. Di dalam tabel ini terdapat variabel c1, c2, dan c3 yang berisi nilai jarak data terhadap setiap pusat klaster. Sedangkan variabel posisi berfungsi untuk menyimpan posisi data terhadap salah satu dari ketiga klaster, dalam hal ini

adalah ke klaster yang terdekat. Tabel Hasil memiliki relasi dengan tabel Data Penduduk, karena di tabel Hasil proses klusterisasi tiap daerah disimpan.

- Tabel Pusat: berfungsi untuk menyimpan nilai rata-rata setiap klaster untuk kelima indikator derajat kesehatan daerah (CBR, CDR, IMR, FMR, dan MMR).
- Tabel Analisis: berfungsi untuk menyimpan hasil analisis pakar untuk setiap klaster. Tabel Analisis memiliki relasi dengan tabel Pusat karena untuk menganalisis pola atau kesamaan yang dimiliki oleh tiap klaster dibutuhkan informasi nilai rata-rata setiap indikator kesehatan.

Rancangan basis data beserta daftar variabel dan relasi yang terbentuk antar tabel dapat dilihat pada Gambar 2.

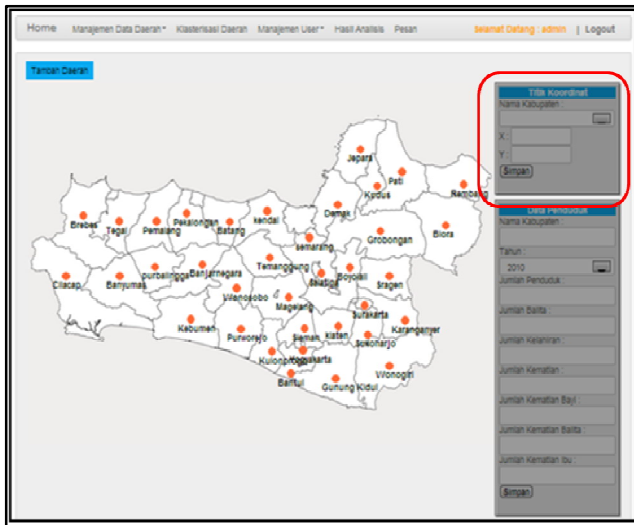


Gambar 2. Relasi Antar Tabel

V. HASIL DAN PEMBAHASAN

Secara umum sistem klusterisasi data indikator kesehatan daerah ini memiliki tiga antarmuka utama, yaitu:

- Halaman Manajemen Data Daerah
 Pada antarmuka ini terdapat tiga fungsi utama, yaitu manajemen data kabupaten, manajemen data penduduk per kabupaten, dan manajemen peta. Di awal implementasi, Admin memasukkan informasi umum, seperti nama dan gambar kabupaten.
 Selanjutnya Admin memetakan posisi tiap kabupaten di peta. Dalam hal ini untuk menginputkan titik koordinat daerah, Admin cukup mengarahkan kursor pada bagian kosong di peta. Posisi koordinat yang disimpan merupakan posisi koordinat monitor (X dan Y), bukan posisi koordinat sesungguhnya di permukaan bumi, sehingga tidak memerlukan sistem proyeksi ataupun sistem koordinat tertentu. Visualisasi proses penentuan titik koordinat tiap daerah dapat dilihat pada Gambar 3.



Gambar 3. Penentuan Posisi Kabupaten di Peta

b. Halaman Klasterisasi Daerah

Pada bagian klasterisasi daerah, Admin akan memproses data mentah yang telah dimasukkan oleh Staf dengan cara memilih data pada tahun tertentu dan kemudian menyimpan data hasil klasterisasi seperti tampak pada Gambar 4.

Adapun data yang ditampilkan berupa kabupaten yang telah dikelompokkan ke dalam tiga kelas (C1, C2, C3). Perbedaan antar kelas dapat dilihat pada warna point yang melekat pada tiap daerah. Selain itu juga terdapat tabel sederhana yang memperlihatkan distribusi data sebagai hasil klastering akhir dan informasi titik pusat klaster berupa nilai rata-rata tiap indikator. Untuk melihat detail informasi tiap klaster, berupa nama kabupaten dan nilai indikator kesehatannya juga dapat dilakukan dengan cara memilih salah satu klaster, misal dipilih klaster yang berwarna merah, sehingga tampil legenda kesehatan daerah seperti tampak pada Gambar 5.

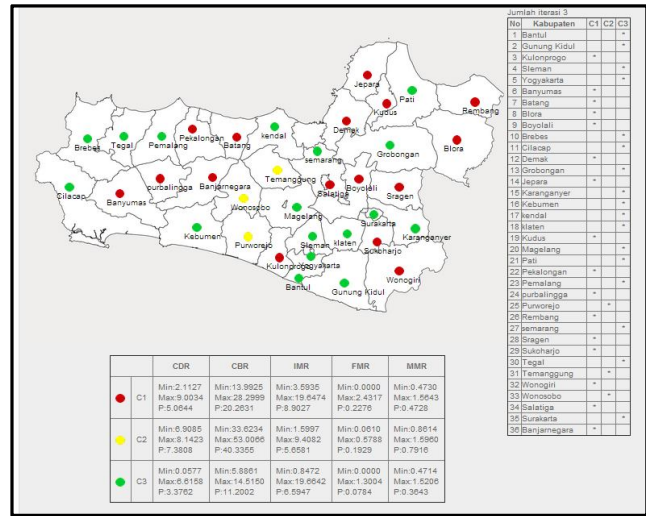
c. Halaman Hasil Analisis Klastering

Setelah data selesai diklaster, Pakar dapat memasukkan analisis terkait hasil klasterisasi data. Untuk memasukkan hasil analisis, Pakar harus memilih tahun dan klaster tertentu sehingga akan tampil data indikator kesehatan sesuai kriteria yang dimasukkan. Tampilan antarmuka halaman analisis klastering dapat dilihat pada Gambar 7.

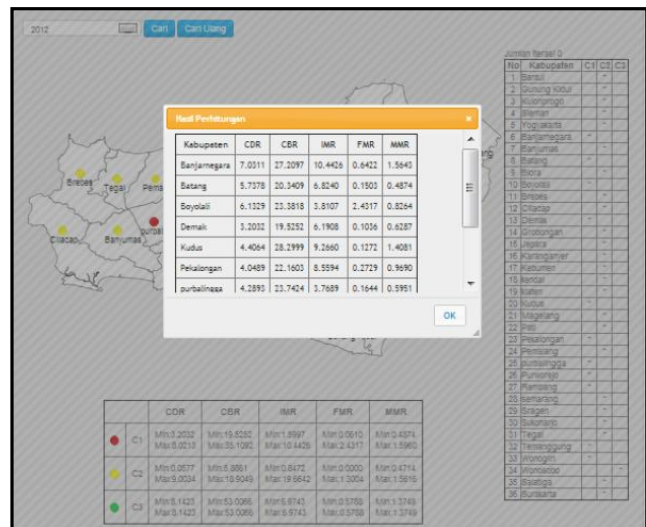
Sebagai contoh, berdasarkan hasil klastering yang tampak pada Gambar 6, untuk indikator CDR diperoleh titik pusat C1=2.59, C2=1.93, dan C3=5.69. Hasil ini dapat dijadikan acuan bagi Pakar untuk membandingkan ketiga klaster. Misal, Pakar menyatakan bahwa CDR untuk daerah-daerah yang masuk di klaster 1 atau C1 adalah Sedang, C2 masuk dalam kategori Rendah, dan CDR untuk C3 termasuk kategori Tinggi. Hal yang sama juga dapat diterapkan untuk keempat indikator mortalitas lainnya.

Berdasarkan hasil perbandingan antar titik pusat antar indikator, maka Pakar dapat menyimpulkan untuk C1 (Klaster

1) memiliki kemiripan dalam hal angka kematian kasar (CDR) sedang, angka kelahiran kasar (CBR) sedang, angka kematian bayi (IMR) rendah, angka kematian balita (FMR) sedang, dan angka kematian ibu (MMR) rendah



Gambar 4. Hasil Klasterisasi Data Indikator Kesehatan



Gambar 5. Keterangan Kesehatan Daerah Sebuah Klaster

	CDR	CBR	IMR	FMR	MMR
C1	Min:0.0577 Max:6.6158 P:2.5879	Min:5.8861 Max:14.5150 P:11.0281	Min:0.8472 Max:19.6642 P:3.1422	Min:0.0000 Max:1.3004 P:0.0867	Min:0.4714 Max:1.5206 P:0.2238
C2	Min:6.9085 Max:8.1423 P:1.9377	Min:27.8673 Max:53.0066 P:5.3120	Min:1.5997 Max:9.4082 P:7.1092	Min:0.0610 Max:0.5788 P:0.0182	Min:0.8614 Max:1.5960 P:0.2851
C3	Min:2.1127 Max:9.0034 P:5.6982	Min:13.9925 Max:28.2999 P:26.3867	Min:3.7689 Max:19.6474 P:6.0594	Min:0.0000 Max:2.4317 P:0.2069	Min:0.4730 Max:1.5643 P:0.5198

Gambar 6. Informasi Hasil Klastering

Home Klasterisasi Daerah Halaman Analisis Ubah Password

2012 C1

Data Cluster

Kabupaten	CDR	CBR	IMR	FMR	MMR
Benjaregara	27.2087	7.0311	10.4426	0.6422	1.5643
Batang	20.3409	5.7378	6.8240	0.1503	0.4874
Bojone	23.3818	6.1329	3.8107	2.4317	0.8264
Cemak	19.5252	3.2032	6.1906	0.1036	0.6287
Kudus	25.2999	4.4064	9.2660	0.1272	1.4081
Pekalongan	22.1603	4.0489	8.5594	0.2729	0.9690
Purbalingga	23.7424	4.2890	3.7689	0.1644	0.5951
Purworejo	35.1092	8.0213	1.5997	0.0610	0.9114
Ramang	25.8363	5.5637	6.2682	0.0829	1.2663
Temanggung	33.6234	6.9085	9.4082	0.3801	1.5960
Wonogri	27.8673	7.9822	3.5635	0.0693	1.1975

Pusat Cluster

CDR	CBR	IMR	FMR	MMR
2.5879	11.0281	3.1422	0.0967	0.2238

Form Analisis

TEXT

Gambar 7. Form Analisis Hasil Klastering

VI. KESIMPULAN DAN SARAN

A. KESIMPULAN

Secara umum, sistem klasterisasi data kesehatan penduduk untuk menentukan rentang derajat kesehatan daerah dengan metode k-Means ini dapat disimpulkan beberapa hal sebagai berikut:

1. Merupakan sebuah aplikasi yang dapat digunakan untuk mengelompokkan daerah kabupaten ke dalam 3 klaster berdasarkan kemiripan karakteristik daerah yang ditinjau dari nilai lima buah indikator mortalitas derajat kesehatan daerah, yaitu angka kelahiran kasar (CBR), angka kematian kasar (CDR), angka kematian bayi (IMR), angka kematian balita (FMR), dan angka kematian ibu (MMR).
2. Menggunakan algoritma k-Means dalam mengelompokkan data berdasarkan metode minimum maksimum untuk normalisasi data dan

metode *Euclidean Distance* untuk menghitung jarak antar data.

3. Memanfaatkan media peta sebagai alat bantu untuk menggambarkan distribusi hasil pengelompokkan pengelompokkan kabupaten berdasarkan kemiripan nilai indikator kesehatan daerah.

B. SARAN

Saran untuk pengembangan aplikasi ini pada penelitian selanjutnya adalah:

1. Perlunya dilakukan analisis terkait jumlah klaster yang optimal terhadap data histori yang terkumpul sebagai salah satu bentuk validasi teknik klastering yang digunakan.
2. Perlunya mengembangkan hasil klastering menggunakan teknik klasifikasi yang dapat dimanfaatkan untuk memprediksi output kategori daerah.
3. Perlunya mengembangkan fitur untuk analisis hasil klastering yang lebih memudahkan kerja Pakar dalam menganalisis hasil klastering.
4. Perlunya mengembangkan fitur untuk mengeksport data ke format dokumen yang memudahkan berbagi informasi antar pihak yang berkepentingan.

DAFTAR PUSTAKA

- [1] D. Kesehatan, "24 Indikator Kesehatan Dalam IPKM," 1 April 2012. [Online]. Available: <http://www.depkes.go.id/index.php?vw=2&id=1337>. [Accessed 1 April 2014].
- [2] D. Kesehatan RI, "Indeks Pembangunan Kesehatan Masyarakat," Kementrian Kesehatan RI, Jakarta, 2010.
- [3] P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Boston, USA: Pearson International Edition, 2005.
- [4] D. Larose, Introduction to Data Mining, Wiley and Sons, 2005.
- [5] S. Kusumadewi, "Sistem Fuzzy untuk Klasifikasi Indikator Kesehatan Daerah," in *Teknoin*, Yogyakarta, 2007.
- [6] J. Han, M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2000.
- [7] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, Advances in Knowledge Discovery and Data Mining, MIT Press, 1996.

Tabel 1. Data Kondisi Kesehatan Penduduk Provinsi Jawa Tengah dan Daerah Istimewa Yogyakarta

No	Kabupaten	Jumlah Penduduk	Jumlah Balita	Jumlah Kelahiran	Jumlah Kematian	Jumlah Kematian Bayi	Jumlah Kematian Balita	Jumlah Kematian Ibu
1	Banjarnegara	869.284	119.901	23.653	6112	247	77	37
2	Banyumas	1.553.902	124.409	27.047	5918	283	19	26
3	Batang	706.015	73.167	14.361	4051	98	11	7
4	Blora	829.604	70.142	13.920	4657	179	13	16
5	Boyolali	931.537	97.464	21.781	5713	83	237	18
6	Brebes	1.732.719	48.048	10.199	3657	94	24	13
7	Cilacap	1.641.031	66.957	13.575	4078	200	0	16
8	Demak	1.058.938	86.848	20.676	3392	128	9	13
9	Grobongan	1.308.592	75.207	15.246	2622	164	21	10
10	Jepara	1.097.158	88.459	16.912	2318	296	27	8

(Sumber: www.depkes.go.id)**Tabel 2.** Hasil Perhitungan Data Indikator Mortalitas

No	Kabupaten	CDR	CBR	IMR	FMR	MMR
1	Banjarnegara	7.031	27.210	10.443	0.642	1.564
2	Banyumas	3.808	17.406	10.463	0.153	0.961
3	Batang	5.738	20.341	6.824	0.150	0.487
4	Blora	5.614	16.780	12.858	0.185	1.149
5	Boyolali	6.133	23.382	3.811	2.432	0.826
6	Brebes	2.111	5.886	9.217	0.500	1.275
7	Cilacap	2.485	8.272	14.733	0.000	1.179
8	Demak	3.203	19.525	6.191	0.104	0.629
9	Grobongan	2.004	11.651	10.757	0.279	0.656
10	Jepara	2.113	15.414	17.502	0.305	0.473

Tabel 3. Hasil Normalisasi Data

No	Kabupaten	CDR	CBR	IMR	FMR	MMR
1	Banjarnegara	1.000	1.000	0.484	0.264	1.000
2	Banyumas	0.359	0.540	0.486	0.063	0.447
3	Batang	0.743	0.678	0.220	0.062	0.013
4	Blora	0.718	0.511	0.661	0.076	0.620
5	Boyolali	0.821	0.820	0.000	1.000	0.324
6	Brebes	0.021	0.000	0.395	0.205	0.735
7	Cilacap	0.096	0.112	0.798	0.000	0.647
8	Demak	0.239	0.640	0.174	0.043	0.143
9	Grobongan	0.000	0.270	0.507	0.115	0.168
10	Jepara	0.022	0.447	1.000	0.126	0.000

Tabel 4. Obyek Terpilih Sebagai Titik Pusat Klaster

C1	1.000	1.000	0.484	0.264	1.000
C2	0.540	0.345	0.486	0.063	0.447
C3	0.270	0.022	0.507	0.115	0.168

Tabel 5. Hasil Perhitungan Jarak Beserta Penentuan Klaster

No	Kabupaten	C1	C2	C3	KLASTER
1	Banjarnegara	0.000	0.993	1.485	C1
2	Banyumas	0.993	0.000	0.509	C2
3	Batang	1.121	0.657	0.887	C2
4	Blora	0.730	0.443	0.874	C2
5	Boyolali	1.140	1.196	1.415	C1
6	Brebes	1.443	0.722	0.645	C3
7	Cilacap	1.391	0.630	0.596	C3
8	Demak	1.271	0.464	0.542	C2
9	Grobongan	1.485	0.509	0.000	C3
10	Jepara	1.609	0.772	0.550	C3

Tabel 6. Pusat Klaster Baru

C1	0.910	0.909	0.242	0.632	0.662
C2	0.592	0.504	0.385	0.061	0.306
C3	0.207	0.025	0.675	0.111	0.387

Tabel 7. Iterasi Kedua Perhitungan Jarak dan Penentuan Klaster

No	Kabupaten	C1	C2	C3	KLASTER
1	Banjarnegara	0.570	0.972	1.420	C1
2	Banyumas	0.940	0.241	0.505	C2
3	Batang	0.911	0.418	1.039	C2
4	Blora	0.827	0.474	0.788	C2
5	Boyolali	0.570	1.087	1.501	C1
6	Brebes	1.366	0.900	0.502	C3
7	Cilacap	1.428	0.839	0.326	C3
8	Demak	1.080	0.392	0.736	C2
9	Grobongan	1.334	0.611	0.283	C3
10	Jepara	1.519	0.866	0.560	C3

Tabel 8. Pusat Klaster Baru yang Stabil

C1	0.910	0.909	0.242	0.632	0.662
C2	0.592	0.504	0.385	0.061	0.306
C3	0.207	0.025	0.675	0.111	0.387