
PENGELOMPOKAN MENGGUNAKAN METODE *SUBTRACTIVE FUZZY C-MEAN* (SFCM), STUDI KASUS DEMAM BERDARAH DI JAWA TIMUR

¹Robert Kurniawan, ²Baiq Nurul Haqiqi

^{1,2} Sekolah Tinggi Ilmu Statistik, Jakarta

Alamat e-mail : ¹ robertk@stis.ac.id

ABSTRAK

Metode pengelompokan yang sering digunakan dalam penelitian adalah *Fuzzy C-Mean Cluster* (FCM). Dalam perkembangannya, FCM dikombinasikan metode *Subtractive Clustering* (SC) sehingga didapatkan *Hybrid Subtractive Fuzzy C-Mean* (SFCM). Metode SFCM memiliki keunggulan dari tingkat kecepatan, dalam hal iterasi, dan menghasilkan partisi data yang lebih stabil dan akurat bila dibandingkan dengan metode sebelumnya. Pada penelitian ini, metode SCFM diaplikasikan dengan 13 variabel dari data demam berdarah. Studi kasus demam berdarah pada penelitian ini dilakukan di Provinsi Jawa Timur. Berdasarkan pengolahan dengan metode SFCM didapat hasil pengelompokan dengan 2 kelompok, 3 kelompok, dan 4 kelompok. Dari 6 indeks validasi untuk mengetahui jumlah pengelompokan yang tepat, menunjukkan bahwa pengelompokan menjadi 2 kelompok memberikan hasil pengelompokan yang lebih bagus dibandingkan dengan pengelompokan yang lainnya. Seluruh kabupaten di Pulau Madura menjadi daerah endemi demam berdarah yang perlu diperhatikan oleh Pemerintah Provinsi Jawa Timur. Dan hal ini senada dengan fakta yang dirilis oleh dinas kesehatan Provinsi Jawa Timur, bahwa beberapa wilayah di Madura menjadi daerah KLB yang memerlukan perhatian serius dalam penanganannya.

Kata Kunci : *Fuzzy C-Mean Cluster* (FCM), *Subtractive Clustering* (SC), *Subtractive Fuzzy C-Mean* (SFCM), Demam Berdarah.

PENDAHULUAN

Pengelompokan (*clustering*) merupakan salah satu teknik yang paling penting dalam *data mining* [1]. Salah satu metode pengelompokan yang paling sering digunakan adalah *Fuzzy C-Means* (FCM). Metode FCM memiliki beberapa kelemahan, antara lain membutuhkan banyaknya kelompok dan matriks keanggotaan kelompok yang ditetapkan sebelumnya [2]. Biasanya, matriks keanggotaan kelompok awal diinisialisasikan secara acak yang

menyebabkan metode FCM memiliki masalah inkonsistensi.

Alternatif metode pengelompokan lainnya yang dapat digunakan jika jumlah kelompok tidak diketahui sebelumnya adalah metode *Subtractive Clustering* (SC). SC memperoleh hasil yang lebih konsisten dibandingkan dengan FCM [3]. Selain itu, SC memiliki kecepatan yang lebih baik dibandingkan FCM, namun SC memiliki akurasi yang lebih rendah dibandingkan dengan FCM [4].

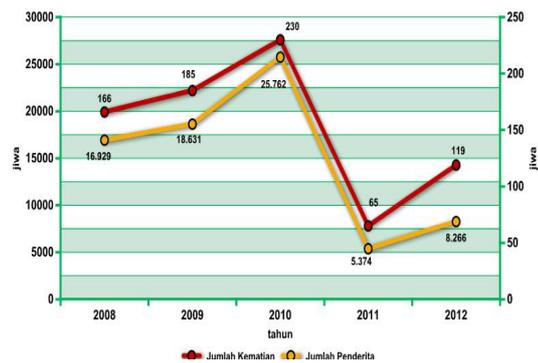
Untuk menjembatani kekurangan dan kelebihan kedua metode tersebut diusulkan sebuah metode baru yang

merupakan penggabungan (*hybrid*) dari keduanya yang disebut *Subtractive Fuzzy C-Means* (SFCM). Metode ini digunakan oleh Liu, Xiao, Wang, Shi, dan Fang [5] dalam penelitiannya dan menyimpulkan bahwa SFCM secara umum memberikan solusi yang lebih baik dibandingkan dengan FCM serta memberikan tingkat kecepatan yang lebih tinggi dalam hal konvergensi fungsi objektif. [4] Kelebihan metode SFCM lainnya dengan menggunakan indeks validitas *Partition Coefficient* (PC) yang menyimpulkan bahwa SFCM dapat meningkatkan kecepatan, mengurangi jumlah iterasi, menghasilkan partisi data yang lebih stabil, dan lebih akurat.

Berdasarkan literatur diatas, dalam penelitian ini metode SFCM cocok diaplikasikan pada data yang berhubungan dengan demam berdarah di Jawa Timur.

Penyakit demam berdarah *dengue* (DBD) atau *Dengue Haemorrhagic Fever* (DHF) merupakan salah satu masalah kesehatan masyarakat yang penting di Indonesia dan mulai dikenal di Indonesia sejak tahun 1968 di Surabaya dan Jakarta. Setelah itu, jumlah kasus DBD terus bertambah seiring dengan semakin meluasnya daerah endemis DBD. Penyakit ini tidak hanya sering menimbulkan Kejadian Luar Biasa (KLB) dengan kematian yang besar tetapi juga menimbulkan dampak buruk sosial dan ekonomi. Kerugian sosial yang terjadi antara lain karena menimbulkan kepanikan keluarga, kematian anggota keluarga, dan berkurangnya usia harapan penduduk.

Penyakit ini pun merupakan salah satu penyakit yang menjadi masalah kesehatan masyarakat dan endemis di hampir seluruh kabupaten/kota di Jawa Timur. DBD juga sudah menjadi masalah yang rutin dihadapi pada setiap musim hujan. Angka kesakitan di Jawa Timur cukup tinggi, meskipun jumlah kematian yang terjadi dapat ditekan [6].



Gambar 1. Trend Penderita DBD dan Jumlah Kematian Akibat DBD Provinsi Jawa Timur Tahun 2008-2012.

Penyakit DBD adalah penyakit infeksi virus akut yang disebabkan oleh virus *Dengue* dan terutama menyerang anak-anak dengan ciri-ciri demam tinggi mendadak dengan manifestasi pendarahan dan bertendensi menimbulkan syok dan kematian. Penyakit ini ditularkan melalui gigitan nyamuk *Aedes Aegypti* dan *Albopictus*. Kedua jenis nyamuk ini terdapat di hampir seluruh pelosok Indonesia kecuali ketinggian lebih dari 1000 meter di atas permukaan laut. Masa inkubasi penyakit ini diperkirakan lebih kurang tujuh hari [7].

Banyak faktor yang berhubungan dengan peningkatan kejadian DBD dan KLB yang sulit atau tidak dapat dikendalikan seperti kepadatan penduduk, mobilitas, lancarnya transportasi (darat, laut, dan udara), perubahan musim dan perubahan iklim dunia, kebersihan lingkungan dan perilaku hidup sehat, serta jenis dan keganasan virusnya [6].

Selanjutnya, tujuan dari penelitian ini adalah untuk mengelompokkan wilayah endemi di Jawa Timur berdasarkan faktor penyebab DBD sehingga pengendalian vektor dapat difokuskan di beberapa wilayah tertentu yang memiliki potensi penyakit DBD tertinggi di Jawa Timur, serta hasil pengelompokan dapat menjadi sasaran sosialisasi terkait

kesadaran untuk berperilaku hidup bersih dan sehat sehingga kematian akibat DBD dapat berkurang.

Untuk mencapai tujuan penelitian tersebut, maka pada penelitian ini digunakan metode SFCM pada data Demam Berdarah. Hasil pengelompokan terbaik dipilih berdasarkan berdasarkan indeks validitas *Partition Coefficient* (PC), *Modified Partition Coefficient* (MPC), *Classification Entropy* (CE), *Partition Index* (PI), *Fukuyama Sugen Index* (FS), dan *Xie Beni Index* (XB).

METODE PENELITIAN

Sumber Data dan Variabel Penelitian

Sumberdata yang digunakan adalah data sekunder dari Dinas Kesehatan Jawa Timur Tahun 2012, dan Badan Pusat Statistik (BPS).

Variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

V_1 = Jumlah penderita demam berdarah di Provinsi Jawa Timur Tahun 2012.

V_2 = Persentase sisi luar/batas dari suatu bangunan atau penyekat dengan bangunan fisik lain yang terbuat dari bambu.

V_3 = Persentase penutup bagian atas suatu bangunan.

V_4 = Persentase bangunan tempat tinggal yang tidak memenuhi syarat kesehatan.

V_5 = Persentase rumah tangga berperilaku tidak sehat.

V_6 = Persentase rumah tangga berperilaku tidak sehat.

V_7 = Persentase rumah tangga tidak bebas jentik.

V_8 = Persentase rumah tangga yang berada di bawah garis kemiskinan.

V_9 = Jumlah kejadian banjir di kabupaten/kota di Jawa Timur Tahun 2012.

V_{10} = Jumlah Rumah Sakit Umum.

V_{11} = Jumlah Dokter Umum di RS.

V_{12} = Proporsi Penduduk Umur kurang dari 15 tahun.

V_{13} = Proporsi Angka Balita Gizi Buruk.

Metode Analisis

Fuzzy C-Means (FCM)

Metode ini ditemukan pertama kali oleh Dunn pada tahun 1973 kemudian dikembangkan lagi oleh Bezdek pada tahun 1981. Ide dasar dari metode ini mirip dengan metode *K-Means*. FCM didasarkan pada logika *fuzzy*, setiap titik data dimasukkan ke suatu kelompok berdasarkan nilai keanggotaannya pada kelompok tersebut.

Algoritma FCM adalah sebagai berikut [8] :

1) Menentukan banyak kelompok (c), *fuzzifier* (m), maksimum iterasi (*MaxIter*), perubahan nilai fungsi objektif terkecil yang diharapkan (ϵ), fungsi objektif awal ($P_0=0$), dan iterasi awal ($t=1$)

2) Membangkitkan bilangan random u_{ik} dengan i merupakan banyak data dan k merupakan banyak kelompok sebagai elemen-elemen awal matriks keanggotaan awal U .

3) Menghitung pusat kelompok ke- i dengan persamaan :

$$p_i = \frac{\sum_{k=1}^N (u_{ik})^m X_k}{\sum_{k=1}^N (u_{ik})^m} \quad (1)$$

dimana u_{ik} adalah nilai keanggotaan objek ke- k dengan pusat kelompok ke- i , X_k adalah objek data ke- k , N adalah banyaknya objek penelitian, dan m adalah *fuzzifier*.

4) Menghitung fungsi objektif pada iterasi ke- t dengan persamaan :

$$J(P, U, X, c, m) = \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m d_{ik}^2(x_k, p_i) \quad (2)$$

dimana c adalah banyak kelompok yang diinginkan, N adalah banyak objek penelitian, u_{ik} adalah nilai keanggotaan objek ke- k pada kelompok ke- i yang merupakan bagian dari matriks U , m adalah *fuzzifier*, dan $d_{ik}^2(x_k, p_i)$ adalah jarak antara vektor pengamatan ke- k dengan pusat kelompok ke- i .

5) Menghitung perubahan matriks keanggotaan dengan persamaan :

$$u_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{ik}^2}{d_{jk}^2}\right)^{\frac{1}{m-1}}} \quad (3)$$

dimana u_{ik} adalah nilai keanggotaan objek ke- k dengan pusat kelompok ke- i , d_{ik}^2 adalah jarak antara objek ke- k dengan pusat kelompok ke- i , d_{jk}^2 adalah jarak antara objek ke- k dengan pusat kelompok ke- j , dan m adalah *fuzzifier*.

6) Cek kondisi berhenti

- Jika $||J_t - J_{t-1}| < \epsilon$ atau $t > MaxIter$ maka berhenti;
- Jika tidak : $t=t+1$, ulangi langkah ke-3

Subtractive Fuzzy C-Means (SFCM)

Subtractive Clustering (SC) merupakan metode pengelompokan yang diperkenalkan oleh Stephen L. Chiu pada tahun 1994. Metode ini merupakan modifikasi dari *Mountain Method* (MM) yang diperkenalkan oleh Yager dan Filev pada tahun 1992. Dimisalkan terdapat n buah titik data $\{x_1, x_2, \dots, x_n\}$ dalam sebuah ruang berdimensi M . Setiap titik data dipertimbangkan sebagai kandidat pusat kelompok yang nilai potensinya dihitung dengan rumus [9] :

$$P_i = \sum_{j=1}^n e^{-\frac{4}{r_\alpha} \cdot \|x_i - x_j\|^2} \quad (4)$$

P_i adalah nilai potensi untuk titik data ke- i atau potensi dari nilai x_i , $\|\cdot\|$ menotasikan jarak Euclidean, dan r_α

adalah sebuah konstanta positif. Konstanta r_α merupakan radius atau jari-jari yang mendefinisikan sebuah lingkungan tetangga. Data yang memiliki potensi yang tinggi adalah data yang memiliki jumlah tetangga paling banyak.

Setelah nilai potensi dari semua titik data dihitung, titik data dengan nilai potensi paling tinggi dipilih sebagai pusat kelompok pertama. Misalkan x_1^* adalah data yang terpilih sebagai pusat kelompok pertama dan P_1^* sebagai nilai potensinya. Nilai potensi untuk setiap titik data diperbaharui dengan rumus[9] :

$$P_i = P_i - P_1^* e^{-\frac{4}{r_\beta} \cdot \|x_i - x_1^*\|^2} \quad (5)$$

dimana r_β merupakan konstanta positif.

Setelah nilai potensi tiap titik data diperbaharui, titik data dengan nilai potensi terbesar dipilih sebagai pusat kelompok kedua. Selanjutnya nilai potensi tiap titik data kembali diperbaharui. Proses ini akan terus berlanjut sampai diperoleh jumlah kelompok yang cukup.

Subtractive Fuzzy C-Means (SFCM) merupakan penggabungan antara metode pengelompokan *Subtractive Clustering* (SC) dan *Fuzzy C-Means* (FCM). SC digunakan untuk menentukan jumlah kelompok dan matriks keanggotaan awal FCM sehingga inialisasi secara random tidak perlu dilakukan. Secara garis besar algoritma metode SFCM adalah sebagai berikut :

- 1) Menghitung potensi awal tiap titik data dengan persamaan (4)
- 2) Menetapkan titik data dengan potensi paling tinggi sebagai pusat kelompok pertama dan P_1^* sebagai nilai potensinya.
- 3) Perbaharui potensi tiap titik data dengan rumus (5)
- 4) Jika rasio potensi titik data tertinggi terbaru dengan potensi titik data tertinggi sebelumnya lebih besar dari *accept ratio* maka titik data dengan potensi tertinggi terbaru ditetapkan

menjadi pusat kelompok dan tahap 3 diulangi kembali. Jika rasio kurang dari *reject ratio* maka iterasi dihentikan berlanjut ke tahap 5.

- 5) Menghitung matriks keanggotaan awal FCM berdasarkan pusat kelompok yang didapatkan melalui metode SC menggunakan persamaan (3). Jika terdapat jarak d_{ik}^2 yang bernilai nol maka u_{ik} (nilai keanggotaan data k pada kelompok ke-i) akan bernilai 1 dan u_{jk} (nilai keanggotaan data k pada kelompok lainnya) akan bernilai 0 [5].
- 6) Melakukan tahapan 3 sampai dengan 6 pada algoritma FCM hingga ditemukannya matriks keanggotaan dan pusat kelompok terakhir.

Indeks Validitas

Beberapa indeks validitas yang sering digunakan dalam penelitian-penelitian adalah :

1) *Partition Coefficient (PC)*

Indeks ini mengukur jumlah *overlapping* antarkelompok. Indeks ini dirumuskan oleh Bezdek sebagai berikut [10] :

$$PC(c) = \frac{1}{N} \sum_{i=1}^c \sum_{k=1}^N u_{ik}^2 \quad (6)$$

dimana N adalah banyak objek penelitian, c adalah banyak kelompok, dan u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i. Indeks ini memiliki rentang $1/c$ sampai 1. Jumlah kelompok yang optimal ditunjukkan oleh nilai PC yang paling besar.

2) *Classification Entropy (CE)*

CE hanya mengukur kekaburan (*fuzziness*) dari partisi kelompok. Indeks ini dirumuskan sebagai berikut [10] :

$$CE(c) = -\frac{1}{N} \sum_{i=1}^c \sum_{k=1}^N u_{ik} \ln(u_{ik}) \quad (7)$$

dimana N adalah banyak objek penelitian, c adalah banyak kelompok,

dan u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i. Indeks ini memiliki rentang 0 sampai $\ln(c)$. Indeks CE yang semakin kecil menunjukkan pengelompokan yang lebih baik.

3) *Partition Index (PI)/Separation and Compactness (SC)*

PI merupakan rasio antara jumlah kepadatan dan pemisahan kelompok-kelompok. Indeks ini dihitung sebagai berikut[11] :

$$PI(c) = \frac{\sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|x_k - v_i\|^2}{N_i \sum_{j=1}^c \|v_j - v_i\|^2} \quad (8)$$

dimana N adalah banyak objek penelitian, N_i adalah banyak objek penelitian kelompok ke-i, c adalah banyak kelompok, u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i, m adalah *fuzzifier*, $\|x_k - v_i\|$ adalah jarak euclidean titik data (x_k) dengan pusat kelompok v_i , dan $\|v_j - v_i\|$ jarak euclidean antar pusat kelompok. Nilai SC yang rendah mengindikasikan partisi kelompok yang lebih baik.

4) *Fukuyama Sugeno Index (FS)*

Indeks ini dirumuskan oleh Fukuyama dan Sugeno sebagai berikut[10] :

$$FS(c) = \frac{\sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|x_k - v_i\|^2}{\sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|v_i - \bar{v}\|^2} \quad (9)$$

dimana N adalah banyak objek penelitian, c adalah banyak kelompok, u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i, m adalah *fuzzifier*, $\|x_k - v_i\|$ adalah jarak euclidean titik data (x_k) dengan pusat kelompok v_i , $\|v_i - \bar{v}\|$ adalah jarak euclidean pusat kelompok v_i dengan rata-rata pusat kelompok. Nilai FS yang

rendah mengindikasikan partisi kelompok yang lebih baik.

5) Xie and Beni's Index (XB)

XB bertujuan untuk menghitung rasio total variasi di dalam kelompok dan pemisahan kelompok. Indeks ini dirumuskan sebagai berikut [10] :

$$XB(c) = \frac{\sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|x_k - v_i\|^2}{N \min_{i,k} \|v_k - v_i\|^2} \quad (10)$$

dimana N adalah banyak objek penelitian, c banyak kelompok, u_{ik} adalah nilai keanggotaan objek ke- k dengan pusat kelompok ke- i . m adalah *fuzzifier*, $\|x_k - v_i\|$ adalah jarak euclidean titik data (x_k) dengan pusat kelompok v_i , dan $\|v_k - v_i\|$ adalah jarak euclidean antar pusat kelompok. Nilai XB yang rendah mengindikasikan partisi kelompok yang lebih baik.

6) Modified Partition Coefficient (MPC)

Indeks ini diajukan oleh Dave (1996) untuk mengatasi kekurangan PC dan CE. Nilai PC dan CE memiliki kecenderungan berubah secara monoton seiring dengan berubahnya nilai c (Wang dan Zhang, 2007). Indeks ini dirumuskan sebagai berikut[10] :

$$MPC(c) = 1 - \frac{c}{c-1} (1 - PC) \quad (11)$$

dimana c adalah banyak kelompok dan PC adalah indeks PC.

Pengolahan data menggunakan aplikasi yang dibangun oleh peneliti dengan *software* R dan berbagai modifikasi.

HASIL PENELITIAN

Dari 13 variabel data Demam Berdarah sebelum dilakukan pengolahan dengan menggunakan metode SFCM, dilakukan standarisasi nilai. Hal ini dilakukan karena untuk menyamakan nilai antar variabel, sehingga lebih memudahkan untuk menganalisisnya.

Pada penelitian ini dicobakan untuk hasil kelompok 2, 3, dan 4. Dimana masing - masing kelompok / *cluster* menggunakan nilai yang sama, yaitu *Squash Factor* (q) = 1,5 ; *Accept Ratio* (ar) = 0,5 ; *Reject Ratio* (rr) = 0,15 ; *Fuzzifier* (m) = 2 dan *Maximum Iteration* ($maxiter$) = 1000. Sedangkan nilai *Radius* (r) untuk masing-masingnya berbeda. Kelompok 2 menggunakan nilai $r = 0,48$; kelompok 3 nilai $r = 0,445$; dan kelompok 4 nilai $r = 0,44$. Perbedaan ini dikarenakan jarak yang digunakan dalam menentukan jumlah kelompok. Jadi, jika nilai jarak (r) diperbesar, maka jumlah kelompok akan semakin kecil, begitu sebaliknya.

Pembentukan kelompok 2, 3 dan 4 menghasilkan indeks validasinya sebagai berikut:

Tabel 1. Indeks Validitas berdasarkan kelompok

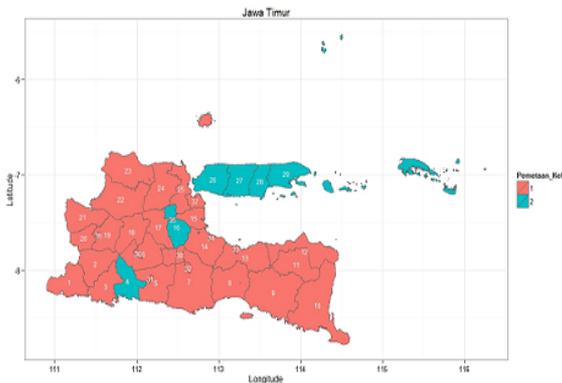
C	PC	MPC	CE	FS	XB	PI
2	0,706	0,413	0,461	105,76	281,68	839,01
3	0,505	0,258	0,818	704,25	237,72	330,85
4	0,428	0,238	104,93	490,56	298,02	156,35

Keterangan:

C = Jumlah Kelompok.

Berdasarkan tabel 1 diatas terlihat bahwa untuk pembentukan menjadi dua kelompok lebih baik jika di bandingkan dengan pembentukan dengan 3 atau 4 kelompok. Dari indeks validitas bisa dilihat bahwa untuk kelompok 2 mempunyai 4 nilai tingkat validitas yang lebih baik dibandingkan kelompok lain. Kelompok 3 dan 4 hanya mempunyai 1 nilai indeks validitas yang baik dibandingkan yang lainnya.

Jika disimulasikan secara visual dengan peta tematik seperti dibawah, bisa terlihat kabupaten/kota mana yang berkelompok membentuk satu kelompok.



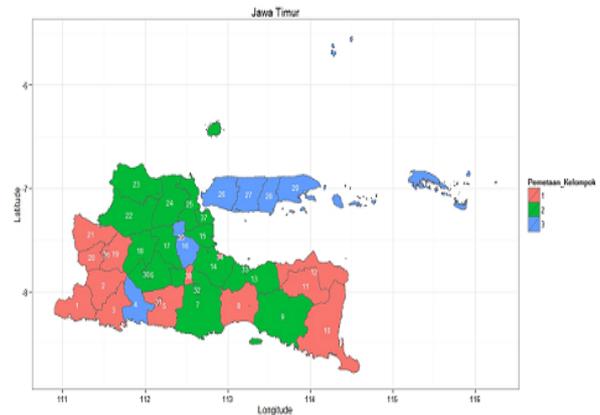
Gambar 2. Peta pembentukan 2 kelompok

Pada gambar 2 terlihat peta persebaran kabupaten/kota berdasarkan data demam berdarah pada pembentukan 2 kelompok sebagai berikut:

Kelompok 1 : Pacitan, Ponorogo, Trenggalek, Blitar, Kediri, Malang, Lumajang, Jember, Banyuwangi, Bondowoso, Situbondo, Probolinggo, Pasuruan, Sidoarjo, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu.

Kelompok 2 : Tulungagung, Mojokerto, Bangkalan, Sampang, Pamekasan, dan Sumenep.

Dari pengelompokan tersebut untuk kelompok 1 bisa disebut dengan kelompok kejadian demam berdarah rendah. Dan kelompok dua bisa disebut kelompok kejadian demam berdarah tinggi. Karena sesuai fakta yang ada, bahwa pada tahun 2012, pemerintah daerah Jawa Timur melalui dinas kesehatan Provinsi Jawa Timur merilis kabupaten/kota dengan Kejadian Luar Biasa (KLB) demam berdarah paling tinggi, yaitu Sumenep, Mojokerto, serta Pamekasan ; dan beberapa kabupaten/kota dengan situasi hampir KLB, yaitu Sampang, Bangkalan dan Tulungagung [12]. Hal tersebut dikarenakan jumlah kematian yang meningkat akibat dari demam berdarah.



Gambar 3. Peta pembentukan 3 kelompok

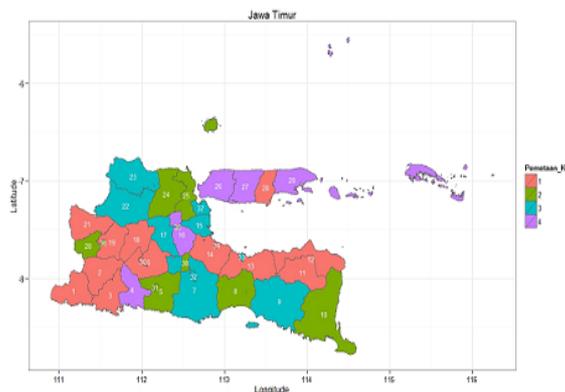
Sedangkan pengelompokan menjadi 3 kelompok berdasarkan gambar 3 diatas terbentuk sebagai berikut:

Kelompok 1 : Pacitan, Ponorogo, Trenggalek, Blitar, Lumajang, Banyuwangi, Bondowoso, Situbondo, Madiun, Magetan, Ngawi, Kota Kediri, Kota Blitar, Kota Malang, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Batu.

Kelompok 2 : Kediri, Malang, Jember, Probolinggo, Pasuruan, Sidoarjo, Jombang, Nganjuk, Bojonegoro, Tuban, Lamongan, Gresik, Kota Probolinggo, Kota Surabaya.

Kelompok 3 : Tulungagung, Mojokerto, Bangkalan, Sampang, Pamekasan, Sumenep.

Hasil pengelompokan menjadi 3 kelompok, untuk kelompok 3 ternyata hasilnya sama dengan pengelompokan dengan 2 kelompok, dimana seluruh kabupaten di Pulau Madura ditambah dengan kabupaten Tulungagung dan Mojokerto menjadi fokus kelompok daerah dengan kejadian demam berdarah paling parah. Hal tersebut dipengaruhi oleh 13 variabel pembentuk yang berhubungan dengan Demam Berdarah.



Gambar 4. Peta pembentukan 4 kelompok

Berdasarkan gambar 4 diatas pengelompokan menjadi 4 kelompok kab/kota yang dipengaruhi oleh 13 variabel demam berdarah sebagai berikut:

Kelompok 1 : Pacitan, Ponorogo, Trenggalek, Kediri, Bondowoso, Situbondo, Probolinggo, Pasuruan, Nganjuk, Madiun, Ngawi, Sampang, Kota Pasuruan.

Kelompok 2 : Blitar, Lumajang, Banyuwangi, Magetan, Lamongan, Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Mojokerto, Kota Madiun, Kota Batu.

Kelompok 3 : Malang, Jember, Sidoarjo, Jombang, Bojonegoro, Tuban, Kota Probolinggo, Kota Surabaya.

Kelompok 4 : Tulungagung, Mojokerto, Bangkalan, Pamekasan, Sumenep.

Berbeda dengan pembentukan kelompok menjadi 2 dan 3, pembentukan kelompok menjadi 4 untuk kelompok yang ke 4 ternyata dari 4 kabupaten di Pulau Madura, ada 1 kabupaten yang keluar dari kelompok, yaitu Kabupaten Sampang. Pindah menjadi kelompok 1.

KESIMPULAN

Berdasarkan hasil penelitian ini , dapat disimpulkan beberapa hal terkait hasil pengelompokan wilayah berdasarkan 13 variabel data demam berdarah. Sesuai dengan nilai validitasnya, maka peneliti menyimpulkan dengan mengelompokkan

menjadi 2 kelompok daerah dengan tingkat kejadian demam berdarah rendah dan tinggi. Sehingga jika di lihat sebaran kabupaten/kota seperti pada gambar 2.

Berdasarkan pengelompokan tersebut, diharapkan Pemerintah Provinsi Jawa Timur lebih memfokuskan untuk pemberantasan dan penurunan angka KLB pada daerah-daerah pada kelompok 2. Karena pada tahun 2015 pun tercatat bahwa yang termasuk dalam kelompok 2, masih terindikasi masuk daerah dengan kondisi KLB DBD.

Beberapa faktor kuat yang menyebabkan kejadian KLB DBD di beberapa daerah di pulau Madura tersebut masih perlu di teliti lebih lanjut. Karena pengelompokan dengan metode SFCM ini hanya berdasarkan radius/jarak antar variabel yang dibangun. Dan belum dapat mengetahui secara pasti variabel mana yang berpengaruh dalam pembentukan kelompok.

DAFTAR PUSTAKA

- [1] Suman dan Mittal, P., *Comparison and Analysis of Various Clustering Methods in Data mining On Education Data Set Using the WEAK Tool*, International Journal of Emerging Trends & Technology in Computer Science, Vol. 3, Issue 2, 2014.
- [2] Le, T. dan Altman, T., *A new initialization method for the Fuzzy C-Means Algorithm using Fuzzy Subtractive Clustering*, Proc. International Conference on Information and Knowledge Engineering, Las Vegas USA, 2011, Vol. 1, pp. 144-150.
- [3] Bataineh, K.M., Naji, M., dan Saqer, M., *A Comparison Study between Various Fuzzy Clustering Algorithms*, Jordan Journal of Mechanical and Industrial

- Engineering, Vol.5, No.4, pp. 335-343, 2011.
- [4] Hossen, J., Rahman, A., Sayeed, S., Samsuddin, K., dan Rokhani, F, *A Modified Hybrid Fuzzy Clustering Algorithm for Data Partitions*, Australian Journal of Basic and Applied Sciences, Vol.5, No.8, pp. 674-681, 2011.
- [5] Liu, W.Y., Xiao, C.J., Wang, B.W., Shi, Y., dan Fang, S.F., *Study on Combining Subtractive Clustering with Fuzzy C-Means Clustering*, International Conference on Machine Learning and Cybernetics, Vol. 5, pp. 2659 – 2662, 2003.
- [6] Dinas Kesehatan Provinsi Jawa Timur. (2013). *Profil Kesehatan Provinsi Jawa Timur 2012*. Surabaya: Dinas Kesehatan Provinsi Jawa Timur.
- [7] Siregar, Faziah A. 2004. *"Epidemiologi dan Pemberantasan Demam Berdarah Dengue (DBD) di Indonesia"*. Fakultas Kesehatan Masyarakat. Universitas Sumatra Utara. [Online]: <http://repository.usu.ac.id/bitstream/123456789/3673/3/fkm-fazidah3.pdf.txt>
- [8] Kusumadewi, S. dan Purnomo, H., *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*, Graha Ilmu, Yogyakarta, 2010.
- [9] Chiu, S.L., *Fuzzy Model Identification Based on Cluster Estimation*, Journal of Intelligent and Fuzzy Systems, Vol. 2, pp. 267-278, 1994.
- [10] Wang, W. dan Zhang, Y., *On fuzzy cluster validity indices*, Fuzzy Sets System, Vol. 158, No. 19, pp.2095-2117, 2007.
- [11] Balasko,B., Abonyi,J. dan Feil,B., *Fuzzy Clustering and Data Analysis Toolbox (for Use With Matlab)*, [online] Available: <http://www.fmt.vein.hu/softcomp/fclusttoolbox/>
- [12] Berita Online Kominfo Pemerintah Provinsi Jawa Timur. Kamis, 26 April 2012 | 11:43. (Diakses tanggal 03 Agustus 2015). [Online]: <http://kominfo.jatimprov.go.id/watch/30866>