

Exploratory Data Analysis dengan JST - Kohonen SOM: Struktur Tingkat Kesejahteraan Daerah Tk II se Jawa Timur

M. Isa Irawan

Program Magister Manajemen Teknologi

Institut Teknologi Sepuluh Nopember

Surabaya

e-mail: mii@its.ac.id

Abstrak:

Sebagai salah satu sumber dana dalam pembiayaan pembangunan daerah, Pendapatan Daerah Sendiri (PDS) memegang peranan penting dalam pelaksanaan otonomi daerah tingkat II. Besarnya PDS menggambarkan potensi pendapatan asli daerah (PAD) dan pajak bumi dan bangunan (PBB). Selain itu untuk menunjang pembangunan daerah dalam rangka pemerataan pembangunan antar daerah pemerintah mengalokasikan sejumlah dana berupa subsidi daerah otonom dan sumbangan pembangunan. Hal ini dilakukan cuma kepada daerah yang pendapatan daerahnya rendah dan kurang berpotensi, sehingga akan mengurangi tingkat kesenjangan antar daerah. Akan tetapi fenomena yang terjadi adalah adanya ketidaksesuaian dalam pengalokasian dana subsidi daerah otonom dan sumbangan pembangunan, akibat dari pengalokasian dana berdasarkan jumlah penduduk.

Pada paper ini dibahas aplikasi jaringan syaraf tiruan Kohonen SOM untuk pengelompokan atau clustering berdasarkan kesamaan atribut yang dimiliki bersama. Banyaknya atribut yang sama akan menggambarkan kedekatan kelompok daerah.

Sistem ternyata mampu menghasilkan kelompok daerah yang terbentuk dari kesamaan dan kemiripan atribut, sehingga dapat diketahui kesesuaian pengalokasian subsidi dan sumbangan pembangunan untuk setiap daerah tingkat II.

Kata kunci : EDA, JST, Kohonen - SOM, Kesejahteraan, PDS

1. Pendahuluan

Prediksi ekonomi secara time series dengan aplikasi jaringan syaraf tiruan (JST) adalah suatu kajian yang sudah banyak dibahas dalam berbagai media ilmiah. Kebanyakan menggunakan JST multi layer perceptron dengan algoritma pembelajaran Backpropagation [2], [3]. Suatu jaringan secara sederhana memanfaatkan data history time series dalam pembelajaran, dan selanjutnya jaringan dapat digunakan untuk memprediksi output yang akan datang.

Dalam suatu komputasi yang lebih kompleks, sangat tidak mungkin untuk

membentuk suatu model karena kurang mencukupinya pengetahuan eksplisit tentang faktor-faktor yang mempengaruhi time-series, atau pengetahuan tersebut perlu cost besar dalam memperolehnya. Dalam keadaan seperti ini metode *exploratory data analysis* (EDA) masih tetap dapat digunakan.

Metode EDA dapat digunakan untuk mengilustrasikan struktur dalam sehimpunan data statistik. Tujuan utama dari EDA adalah menampilkan sehimpunan data statistik agar mudah

dipahami tetapi pada waktu yang bersamaan menyediakan informasi se esensial mungkin dari himpunan data aslinya. Metode EDA adalah perangkat *general purpose* yang dapat mengilustrasikan fitur-fitur esensial dari sehimpunan data, seperti struktur *clustering* dan relasi antara item-item data.

Seseorang bisa melakukan dengan dua tools EDA yang berbeda. Pertama, tool proyeksi Sammon [8] dari himpunan data multi dimensi ke, misalnya bidang dua-dimensi, sambil tetap memungkinkan untuk menampilkan struktur keseluruhan (jarak antara item-item data). Kedua, mendapatkan *cluster* dari data, dimana dalam memanfaatkan data set besar hanya memperhatikan sejumlah kecil cluster saja.

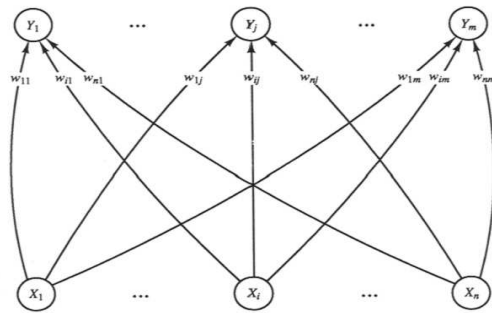
Cluster analysis biasanya didasarkan kepada jarak *ultrametric* yang didefinisikan sepanjang graph hirarki clustering melalui perbandingan langsung dari similaritas setiap dua data set yang terlihat membingungkan.

Algoritma Kohonen-SOM adalah suatu metode yang unik dapat mengkombinasikan proyeksi dan algoritma clustering. Pada saat yang sama Kohonen SOM dapat digunakan untuk memvisualisasi cluster dalam sehimpunan data set dan untuk merepresentasikan data set dalam pemetaan dua-dimensi dengan tetap menjaga kedidaklinieran item-item data.

2. Kohonen Self Organizing Map (SOM)

Jaringan Kohonen SOM (*Self-Organizing Map*) merupakan salah satu model jaringan syaraf yang menggunakan metode pembelajaran *unsupervised* [4].

Jaringan Kohonen SOM terdiri dari dua lapisan (*layer*), yaitu lapisan *input* dan lapisan *output*. Setiap neuron dalam lapisan *input* terhubung dengan setiap neuron pada lapisan *output*. Setiap neuron dalam lapisan *output* merepresentasikan kelas dari *input* yang diberikan. Gambar struktur JST Kohonen terlihat seperti pada Gambar 1.



Gambar1. Struktur JST Kohonen SOM

Penulisan istilah :

X : vektor input pembelajaran.

$$X = (x_1, x_2, \dots, x_j, \dots, x_n).$$

α : *learning rate*

R : *radius neighborhood*

X_i : neuron/node input.

w_{0j} : bias pada neuron output ke- j

Y_j : neuron/node output ke- j

C : konstanta

Berikut ini adalah tahapan dalam algoritma pembelajaran JST Kohonen SOM [1]:

Langkah 0 : Inisialisasi pembobotan w_{ij} dengan nilai random. Menetap parameter learning rate (α), dan radius tetangga (R)

Langkah 1 : Apabila kondisi selesai belum terpenuhi, lakukan langkah 2-8

Langkah 2 : Untuk tiap vektor input $x(x_i, i = 1, \dots, n)$, lakukan langkah 3-5

Langkah 3 : Untuk tiap j ($j = 1, \dots, m$), hitung jarak Euclidean

$$D(j) = \sum_i (w_{ij} - x_i)^2 \quad (1)$$

Langkah 4 : Mencari indeks j dengan jarak $D(j)$ terdekat (minimum)

Langkah 5: Melakukan perbaikan nilai w_{ij} dengan nilai tertentu. Yaitu :

$$w_{ij}(\text{new}) = w_{ij}(\text{old}) + \alpha[x_i - w_{ij}(\text{old})] \quad (2)$$

Langkah 6: Melakukan update learning rate.

$$\alpha_{(t+1)} = \alpha_t \quad (3)$$

Langkah 7: Mereduksi radius dari fungsi tetangga pada waktu tertentu (*epoch*).

Langkah 8: Menentukan kondisi STOP.

3. Struktur Tingkat Kesejahteraan Daerah Tingkat II

Salah satu sumber dana dalam pembiayaan pembangunan daerah adalah Pendapatan Daerah Sendiri (PDS) yang terdiri dari Pendapatan Asli Daerah (PAD), bagian pendapatan dari pemerintah dan pinjaman pemerintah daerah. PAD terdiri dari pajak daerah, retribusi daerah, laba usaha daerah dan penerimaan lain-lain.

Besarnya PDS menggambarkan kemampuan daerah di dalam menggali potensi PAD dan pajak bumi dan bangunan (PBB) menuju kemandirian daerah di bidang pembiayaan tugas-tugas otonomi daerah tersebut. Sementara itu terdapat perbedaan besarnya PDS antar daerah tingkat II. Hal ini menunjukkan pula adanya perbedaan potensi PAD dan PBB untuk masing-masing daerah tersebut.

Selain itu untuk menunjang pelaksanaan otonomi daerah dan dalam upaya mewujudkan pemerataan pembangunan antar daerah, pemerintah mengalokasikan sejumlah dana berupa sumbangan dan bantuan utamanya kepada daerah yang berpendapatan rendah dan kurang potensi, sehingga kesenjangan antar daerah dapat dikurangi. Namun demikian pengalokasian dana bantuan untuk tujuan tersebut tidaklah mudah. Untuk itu perlu adanya suatu cara yang sistematis yang dapat digunakan untuk mengetahui perbedaan PDS secara objektif, sehingga

pengalokasian dana bantuan dapat dilakukan secara proposional atau dengan menggunakan skala prioritas terhadap daerah-daerah yang kurang potensi.

Dengan cara mengelompokkan daerah-daerah tingkat II berdasarkan atribut-atribut yang ada pada PDS dan jumlah penduduk dalam penelitian ini diharapkan dapat diketahui kesesuaian antara perbedaan PDS dengan pengalokasian dana sumbangan dan bantuan yaitu untuk kelompok daerah yang kurang potensi akan memperoleh dana sumbangan dan bantuan yang lebih besar dibandingkan kelompok daerah yang lain. Demikian berlaku sebaliknya kelompok daerah yang berpotensi akan mendapat dana sumbangan dan bantuan yang lebih kecil.

3.1. Data Set

Data yang digunakan dalam penelitian ini berupa data realisasi penerimaan keuangan daerah tingkat II propinsi Jawa Timur tahun anggaran 1999/2000 dan jumlah penduduk pertengahan tahun 2000 per Kabupaten/Kota [5], [6], [7]. Data-data tersebut diperoleh dari Kantor Statistik Propinsi Jawa Timur.

Kecuali Kota Surabaya, daerah tingkat II yang merupakan obyek dari penelitian ini terdiri dari :

A. Daerah Tingkat II Kota :

1. Madiun
2. Mojokerto
3. Pasuruan
4. Probolinggo
5. Malang
6. Blitar
7. Kediri

B. Daerah Tingkat II Kabupaten :

1. Pacitan
2. Ponorogo
3. Trenggalek

4. Tulungagung
5. Blitar
6. Kediri
7. Malang
8. Lumajang
9. Jember
10. Banyuwangi
11. Bondowoso
12. Situbondo
13. Probolinggo
14. Pasuruan
15. Sidoarjo
16. Mojokerto
17. Jombang
18. Nganjuk
19. Madiun
20. Magetan
21. Ngawi
22. Bojonegoro
23. Tuban
24. Lamongan
25. Gresik
26. Bangkalan
27. Sampang
28. Pamekasan
29. Sumenep

Sedangkan atribut-atribut yang digunakan dalam penelitian ini terdiri dari :

A. X_1 = Jumlah Penduduk

B. Pendapatan Daerah Sendiri (PDS)

yang terdiri dari :

1. Pendapatan Asli Daerah (PAD) yang berupa :
 - a. X_2 = Pajak Daerah
 - b. X_3 = Retribusi Daerah
 - c. X_4 = Bagian Laba Usaha Daerah
 - d. X_5 = Penerimaan lain-lain
2. Bagian Pendapatan Dari Pemerintah dan atau Instansi Yang Lebih Tinggi yang berupa :
 - a. X_6 = Bagi Hasil Pajak
 - b. X_7 = Bagi Hasil Bukan Pajak
 - c. X_8 = Subsidi Daerah Otonom
 - d. X_9 = Bantuan Pembangunan
 - e. X_{10} = Penerimaan Lainnya

3. Pinjaman Pemerintah Daerah
 - a. X_{11} = Pinjaman Pemerintah Pusat
 - b. X_{12} = Pinjaman Lembaga Keuangan Dalam Negeri
 - c. X_{13} = Pinjaman Dari Luar Negeri

Data set yang digunakan sebagai input dinormalkan dengan nilai rata-rata sebagai acuan yang analog dengan :

$$f(x) = \begin{cases} 1, & x_i \geq x_{rata-rata} \\ 0, & x_i < x_{rata-rata} \end{cases} \quad (4)$$

Berdasarkan data maka terlihat untuk masing-masing atribut memiliki nilai terendah, nilai tertinggi dan nilai rata-rata sebagai berikut :

Tabel 1.
Deskriptif data penerimaan daerah

Atribut	Minimum	Maksimum	Rata-rata
X_1	108.045	2.382.258	876.547
X_2	413.372	18.910.540	3.638.099
X_3	1.390.530	18.912.361	3.804.806
X_4	0	2.262.638	194.017
X_5	147.713	8.306.518	1.302.204
X_6	3.839.111	26.197.746	9.973.031
X_7	19.415	373.645	75.521
X_8	14.847.087	125.556.506	60.269.198
X_9	8.159.897	46.938.060	19.380.742
X_{10}	0	859.237	95.291
X_{11}	0	1.673.204	69.384
X_{12}	0	1.248.750	34.687
X_{13}	0	0	0

Selanjutnya dari nilai rata-rata seperti Tabel 3.2 maka akan dijadikan acuan untuk menentukan input dari data menuju ke input pada JST *Kohonen* dengan pengkodean 1 dan 0 pada Tabel 2.

Kemudian data-data input tersebut diproses oleh sistem JST

sehingga menghasilkan output berupa kelompok daerah tingkat II berdasarkan penerimaan daerah.

3.2. Komputasi JST Kohonen SOM

Jaringan Kohonen-SOM yang digunakan digambarkan pada Gambar 1. Neuron *input* dihubungkan dengan neuron *output* dengan koneksi bobot, yang mana bobot ini selalu diperbaiki pada proses iterasi pelatihan jaringan.

Aliran informasi sistem JST untuk pengelompokan tingkat kesejahteraan daerah tingkat II ini diawali dengan dimasukkannya data penerimaan daerah. Data-data inilah yang akan berfungsi sebagai data *input* awal, selain parameter *input*, berupa *learning rate* (α), dan radius *neighborhood* (R)

4. Hasil

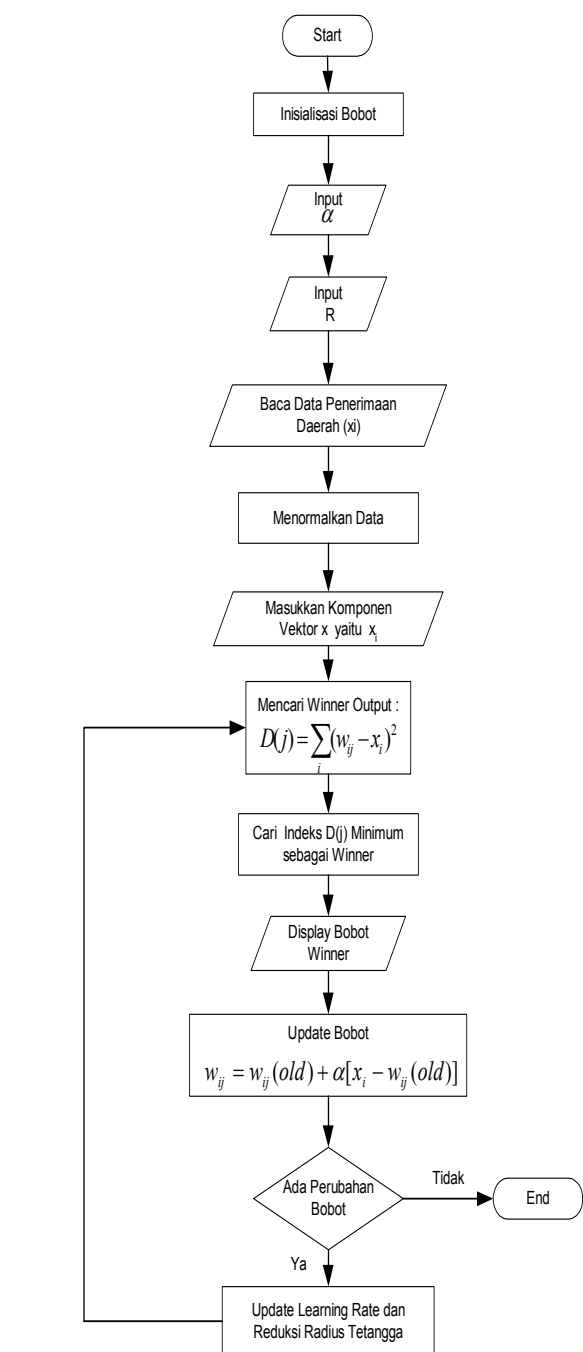
4.1. Uji Coba Perangkat lunak

Berdasarkan hasil uji coba perangkat lunak daerah tingkat II terjadi kemiripan antar satu dengan yang lain yang bergabung membentuk empat kelompok.

Kelompok pertama terdiri dari 15 daerah tkt II yaitu: Kota. Madiun, Kab. Pacitan, Kab. Trenggalek, Kab. Bondowoso, Kab. Situbondo, Kab. Mojokerto, Kab. Madiun, Kab. Ngawi, Kab. Bangkalan, Kab. Sampang, Kota Kediri, Kota Blitar, Kota Probolinggo, Kota Pasuruan, dan Kota Mojokerto.

Kelompok kedua terdiri dari 11 daerah tingkat II yaitu Kab. Ponorogo, Kab. Tulungagung, Kab. Blitar, Kab. Lumajang, Kab. Probolinggo, Kab. Jombang, Kab. Nganjuk, Kab. Magetan, Kab. Bojonegoro, Kab. Lamongan, Kab. Pamekasan.

Kelompok ketiga terdiri dari 7 daerah tingkat II yaitu Kab. Kediri, Kab. Malang, Kab. Jember, Kab. Banyuwangi, Kab. Pasuruan, Kab. Tuban, dan Kab. Sumenep.



Gambar 3. Diagram Alir JST Kohonen

Sedangkan kelompok keempat terdiri dari 3 daerah tingkat II yaitu Kab. Sidoarjo, Kab. Gresik, dan Kota Malang.

Dari data terungkap bahwa untuk kedelapan atribut, nilai rata-rata terendah terletak pada kelompok I, sedangkan nilai rata-rata tertinggi pada kelompok IV. Kelompok II dan III berada di antara kelompok I dan IV.

Dengan demikian dapat dikatakan bahwa :

1. Kelompok I merupakan kelompok daerah tingkat II dengan pendapatan daerah sendiri dan jumlah penduduk yang kecil.
2. Kelompok II merupakan kelompok daerah tingkat II dengan pendapatan daerah sendiri dan jumlah penduduk sedang.
3. Kelompok III merupakan kelompok daerah tingkat II dengan pendapatan daerah sendiri dan jumlah penduduk besar.
4. Kelompok IV merupakan daerah tingkat II dengan pendapatan daerah sendiri dengan jumlah yang paling besar dan jumlah penduduk kecil. Daerah ini merupakan daerah yang paling optimal untuk tingkat kesejahteraannya. Karena dengan jumlah penduduk yang kecil, jumlah pendapatan daerah sendirinya besar.

4.2. Perbedaan Sumbangan dan Bantuan Antar Kelompok Daerah Tingkat II

Berdasarkan hasil pengelompokan diperoleh 4 kelompok daerah tingkat II dan melihat dari hasil rata-rata penerimaan daerah sendiri dan jumlah penduduk menunjukkan perbedaan yang nyata. Untuk mengetahui adanya perbedaan keempat kelompok dalam hal ini penerimaan subsidi daerah otonom dan sumbangan pembangunan dilakukan suatu analisa terhadap rata-rata dari sumbangan dan bantuan dari daerah tingkat II selama ini.

Jika dilihat dari data BPS yang ada alokasi bantuan subsidi daerah otonom dan sumbangan pembangunan sebagai penerimaan daerah tingkat II untuk keempat

kelompok daerah tingkat II yang telah dibentuk di atas menunjukkan adanya perbedaan nyata. Kelompok I merupakan kelompok dengan alokasi bantuan subsidi otonom dan sumbangan pembangunan yang terkecil dan sebaliknya kelompok III merupakan kelompok dengan alokasi bantuan subsidi otonom dan sumbangan pembangunan yang paling besar.

Alokasi subsidi daerah otonom dan sumbangan pembangunan didasarkan jumlah penduduk. Daerah tingkat II yang berpenduduk kecil akan memperoleh subsidi daerah otonom dan sumbangan pembangunan yang kecil, sebaliknya daerah tingkat II yang berpenduduk besar memperoleh subsidi daerah otonom dan sumbangan pembangunan yang besar pula. Sementara itu dari hasil pengelompokan menunjukkan bahwa kelompok I yang merupakan kelompok dengan jumlah penduduk dan pendapatan daerah sendiri yang kecil ternyata memperoleh subsidi daerah otonom dan sumbangan pembangunan yang kecil dan sebaliknya untuk kelompok III memperoleh subsidi daerah otonom dan sumbangan pembangunan yang besar.

Jika memperhatikan tujuan diadakannya subsidi daerah otonom dan sumbangan pembangunan yaitu untuk mewujudkan pemerataan antar daerah guna mengurangi kesenjangan antar daerah tingkat II, maka untuk daerah tingkat II dengan pendapatan daerah yang kecil seharusnya mendapatkan subsidi daerah otonom dan sumbangan pembangunan yang besar.

Dari uraian di atas terkesan bahwa pengalokasian subsidi daerah otonom dan sumbangan pembangunan

didasarkan pada jumlah penduduk dan kurang memperhatikan pendapatan daerah sendiri yang merupakan potensi dari keuangan daerah.

5. Penutup

Melihat fenomena yang ada, untuk mengalokasikan subsidi daerah otonom dan sumbangan pembangunan yang sesuai dengan tujuannya tidaklah mudah karena membutuhkan suatu cara yang benar-benar dapat digunakan untuk tujuan tersebut.

Dari hasil pengelompokan berdasarkan jumlah penduduk dan penerimaan daerah sendiri yang telah dilakukan, sedikit banyak dapat memberi gambaran terhadap daerah-daerah tingkat II yang kurang potensi, sehingga dengan skala prioritas dan proposional pengalokasian subsidi daerah otonom dan sumbangan pembangunan dapat dilaksanakan sesuai dengan tujuannya.

Daftar Pustaka

- [1] Fauset, L, (1994), *Fundamental of Neural Networks, Architectures, Algorithms,, and Aplications*, Prentice Hall, New Jersey
- [2] Irawan, M. I., (1998), "Multivariate Time Series Prediction of Crude Oil Prices based on ANNs", *Proceeding of International Congress of Soft Computing EUFIT 1998*, Aachen, Germany.
- [3] Irawan, M.I. , (1998), "Data Mining Application for Financial Credit Scoring based on ANNs", *Proceeding of ISSM*, Paderborn, Germany
- [4] Kaski, S., (1997), *Acta Politecnica Scandinavica: Data Exploration using Self Organizing Maps*, Helsinki University of Technology, Finland
- [5] Noname, (2000), *Statistik Keuangan Pemerintah Daerah Kabupaten/*

Kotamadya se Jawa Timur, Biro Pusat Statistik Jawa Timur

- [6] Noname, (2001), *Sensus Kependudukan 2000 Propinsi Jawa Timur*, Biro Pusat Statistik Jawa Timur
- [7] Noname, (2001), *JawaTimur Dalam Angka th 2000*, Biro Pusat Statistik Jawa Timur.
- [8] Sammon, J.W., (1969), *IEEE Tr. Computers*, C-8.

Tabel2 Hasil Pengkodean Data

Nama_daerah	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃
Kdy. Madiun	0	0	0	0	0	0	0	1	1	0	1	1	0
Pacitan	0	0	0	0	0	0	0	1	0	1	1	1	0
Ponorogo	1	0	0	0	0	0	0	0	1	1	1	0	0
Trenggalek	0	0	0	1	0	1	1	1	1	0	1	1	0
Tulungagung	1	0	1	0	0	0	0	0	1	0	1	1	0
Blitar	1	0	0	0	1	0	0	0	0	0	1	1	0
Kediri	1	1	1	0	0	1	0	0	0	0	1	1	0
Malang	1	1	1	0	0	1	0	0	0	1	1	1	0
Lumajang	1	0	1	0	0	0	0	1	1	0	1	1	0
Jember	1	1	1	1	0	1	0	0	0	0	0	1	0
Banyuwangi	1	1	1	0	0	1	0	0	0	1	0	1	0
Bondowoso	0	0	0	0	0	0	0	1	1	1	1	1	0
Situbondo	0	0	0	0	0	0	0	1	0	0	1	1	0
Probolinggo	1	0	0	0	0	0	0	0	1	0	1	1	0
Pasuruan	1	1	1	0	0	1	0	0	0	0	1	1	0
Sidoarjo	1	1	1	1	1	1	0	0	1	0	1	1	0
Mojokerto	1	1	0	0	0	0	0	1	1	0	1	1	0
Jombang	1	0	0	0	1	0	0	0	1	0	1	1	0
Nganjuk	1	0	1	0	0	0	0	0	1	1	1	1	0
Madiun	0	0	0	0	0	0	0	0	1	0	1	1	0
Magetan	0	0	1	0	1	0	0	0	1	0	1	1	0
Ngawi	0	0	0	0	0	0	0	0	1	0	1	1	0
Bojonegoro	1	0	1	0	1	0	0	0	1	1	1	1	0
Tuban	1	1	1	0	0	1	0	0	0	0	1	1	0
Lamongan	1	0	0	0	0	0	0	0	0	0	1	1	0
Gresik	1	1	0	0	1	1	0	1	1	0	1	1	0
Bangkalan	0	0	1	0	0	0	0	1	1	0	0	1	0
Sampang	0	0	0	0	0	0	0	1	1	0	1	1	0
Pamekasan	0	0	1	0	0	0	0	0	1	0	1	1	0
Sumenep	1	0	0	0	0	1	0	0	0	0	1	1	0
Kdy. Kediri	0	0	0	0	1	0	0	1	1	0	1	1	0
Kdy. Blitar	0	0	0	0	0	0	0	1	1	0	1	1	0
Kdy. Malang	0	1	1	1	1	1	0	1	1	1	1	1	0
Kdy. Probolinggo	0	0	0	0	0	0	0	1	1	0	1	1	0
Kdy. Pasuruan	0	0	0	0	0	0	0	1	1	0	1	1	0
Kdy. Mojokerto	0	0	0	0	0	0	0	1	1	0	1	1	0