

---

---

## METODE DBSCAN UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI JAWA TENGAH BERDASARKAN PRODUKSI PADI SAWAH DAN PADI LADANG

<sup>1</sup>Diah Safitri, <sup>2</sup>Triastuti Wuryandari, <sup>3</sup>Rita Rahmawati

<sup>1,2,3</sup>Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro  
Email : diahsafitri.fifi@gmail.com

### ABSTRAK

Padi merupakan tanaman yang penting di Jawa Tengah karena nasi merupakan makanan pokok sebagian besar masyarakat di Jawa Tengah. Provinsi Jawa Tengah terdiri dari 35 kabupaten/kota, dalam penelitian ini berdasarkan produksi padi sawah dan padi ladang, kabupaten/kota di Provinsi Jawa Tengah akan dikelompokkan menjadi beberapa kelompok menggunakan metode *Density Based Spatial Clustering Algorithm With Noise* (DBSCAN) berdasarkan produksi padi sawah dan padi ladang, dimana pada masing-masing kelompok dapat dilihat karakteristiknya mengenai potensi produksi padi sawah dan padi ladang. Metode DBSCAN adalah metode yang tangguh untuk mendeteksi *noise*. Kelompok 1 terdiri dari Kabupaten/Kota di Provinsi Jawa Tengah selain yang sudah masuk dalam kelompok 2 dan 3, kelompok 1 adalah kelompok yang mempunyai hasil produksi padi sawah terendah dibandingkan dengan kelompok yang lain. Kelompok 2 adalah kelompok yang mempunyai hasil produksi padi ladang tertinggi dibandingkan dengan kelompok yang lain. Kelompok 2 terdiri dari Kabupaten Kebumen dan Kabupaten Blora. Kelompok 3 terdiri dari Kabupaten Sragen, Kabupaten Grobogan, Kabupaten Pati, Kabupaten Demak, dan Kabupaten Brebes, adalah kelompok yang mempunyai hasil produksi padi sawah tertinggi dibandingkan dengan kelompok yang lain. Pada penelitian ini ditemukan 2 *noise*, yaitu Kabupaten Cilacap dan Kabupaten Wonogiri.

**Kata Kunci:** DBSCAN, Padi Sawah, Padi Ladang

### PENDAHULUAN

Padi terdiri dari padi sawah dan padi ladang. Padi sawah adalah padi yang ditanam di lahan sawah, dimana lahan sawah adalah lahan pertanian yang berpetak-petak dan dibatasi oleh pematang (galengan), saluran untuk menahan/menyalurkan air, yang biasanya ditanami padi sawah tanpa memandang dari mana diperolehnya atau status lahan tersebut. Lahan sawah mencakup sawah pengairan, tadah hujan, sawah pasang surut, rembesan, lebak dan lain sebagainya. Padi ladang adalah padi yang

ditanam di tegal/kebun/ladang atau huma. Luas bersih adalah luas sawah secara keseluruhan (luas kotor) dikurangi dengan luas pematang/galengan dan luas saluran air. Luas panen adalah luasan tanaman yang dipungut hasilnya setelah tanaman tersebut cukup umur. Bentuk produksi dari padi berupa gabah kering giling [2]. Padi merupakan tanaman yang penting di Jawa Tengah karena nasi merupakan makanan pokok sebagian besar masyarakat di Jawa Tengah.

Provinsi Jawa Tengah terdiri dari 35 kabupaten/kota, dalam penelitian ini

berdasarkan produksi padi sawah dan padi ladang, kabupaten/kota di Provinsi Jawa Tengah akan dikelompokkan menjadi beberapa kelompok berdasarkan produksi padi sawah dan padi ladang. Dalam penelitian ini, 35 kabupaten/kota di Jawa Tengah dikelompokkan menjadi beberapa kelompok berdasarkan produksi padi sawah dan padi ladang, dimana pada masing-masing kelompok dapat dilihat karakteristiknya mengenai potensi produksi padi sawah dan padi ladang, sehingga dapat sebagai masukan bagi pihak terkait dalam mengambil kebijakan guna mendukung ketahanan dan keamanan pangan.

Menurut strukturnya, pengelompokan dibagi menjadi dua yaitu pengelompokan hierarki dan sekatan. Menurut keanggotaan data dalam kelompok, pengelompokan dibagi menjadi dua, yaitu eksklusif dan tumpang tindih. Menurut kategori kekompakan, pengelompokan terbagi menjadi dua yaitu komplet dan parsial. Metode *Density Based Spatial Clustering Algorithm With Noise* (DBSCAN) menurut strukturnya, masuk dalam kategori metode pengelompokan sekatan, menurut keanggotaan data dalam kelompok, masuk dalam kategori eksklusif, dimana sebuah data dipastikan hanya menjadi anggota satu kelompok dan tidak menjadi anggota kelompok yang lain. Metode DBSCAN adalah metode yang tangguh untuk mendeteksi data yang menyimpang, data yang menyimpang disebut *noise (outlier)* [6].

Pengelompokan 35 kabupaten/kota di Provinsi Jawa Tengah pada penelitian ini menggunakan metode DBSCAN. Dalam penelitian ini, 35 kabupaten/kota di Jawa Tengah dikelompokkan menjadi beberapa kelompok berdasarkan produksi padi sawah dan padi ladang, dimana pada masing-masing kelompok dapat dilihat karakteristiknya mengenai potensi produksi padi sawah dan padi ladang, sehingga dapat sebagai masukan bagi

pihak terkait dalam mengambil kebijakan guna mendukung ketahanan dan keamanan pangan.

*Clustering* merupakan proses pengelompokan sejumlah besar data menjadi beberapa kelas sesuai dengan ciri khasnya masing-masing. Algoritma *clustering* yang paling efisien untuk menentukan kelompok pada data dengan kepadatan yang berbeda adalah algoritma *density based clustering* [4].

*Density Based Spatial Clustering Algorithm with Noise* (DBSCAN) adalah algoritma pengelompokan yang didasarkan pada kepadatan (*density*) data. Konsep kepadatan yang dimaksud dalam DBSCAN adalah jumlah data yang berada dalam radius *MinPts* (Jumlah minimal data dalam radius ), data tersebut masuk dalam kategori kepadatan yang diinginkan, jumlah data dalam radius tersebut termasuk data inti itu sendiri. Konsep kepadatan seperti ini melahirkan tiga macam status dari setiap data, yaitu inti (*core*), batas (*border*), dan noise (*noise*). Sebuah data akan dimasukkan sebagai inti jika jumlah data tetangga dan dirinya sendiri pada radius berjumlah *MinPts*. Nilai radius dan *MinPts* ini ditetapkan secara mandiri. Untuk data yang jumlah tetangga dan dirinya sendiri dalam radius kurang dari *MinPts*, tetapi tetangganya menjadi inti karena kehadirannya, data tersebut dikategorikan sebagai batas. Jika jumlah tetangga dan dirinya sendiri dalam radius kurang dari *MinPts* dan tidak ada tetangga yang menjadi inti karena kehadirannya, data tersebut dikategorikan sebagai noise [6].

DBSCAN adalah salah satu contoh pelopor perkembangan teknik pengelompokan berdasarkan kepadatan atau yang biasa dikenal dengan sebutan *density based clustering* [4].

Menurut [5], keuntungan dari DBSCAN :

1. DBSCAN tidak memerlukan informasi tentang banyaknya kelompok yang akan dibentuk.
2. DBSCAN dapat memperoleh kelompok berbentuk sembarang, bahkan dapat menemukan kelompok yang benar-benar dikelilingi oleh ( tetapi tidak terhubung) pada cluster yang berbeda
3. DBSCAN mempunyai informasi tentang noise
4. DBSCAN membutuhkan hanya dua parameter yang sebagian besar tidak sensitif terhadap urutan titik dalam basis data.

Algoritma DBSCAN membutuhkan dua parameter penting, yaitu  $Eps$  ( ) dan jumlah data tetangga minimal untuk membentuk kelompok ( $MinPts$ ). Algoritma dimulai dari sembarang data yang belum dikunjungi. Data ini kemudian dibaca jumlah tetangganya pada radius . Jika jumlahnya mencukupi (lebih dari atau sama dengan ), data akan ditandai sebagai inti dan tetangganya sebagai batas (kecuali yang sudah ditandai sebagai inti), kemudian sebuah kelompok baru akan mulai dibentuk. Jika tidak mencukupi, data tersebut akan ditandai sebagai noise. Tandai data tersebut sebagai data yang sudah dikunjungi. Langkah tersebut dapat dilakukan secara rekursif pada setiap data yang menjadi tetangganya dan belum dikunjungi [6].

Setelah terbentuk kelompok-kelompok, dihitung koefisien silhouette. Di dalam [1] dikatakan bahwa hasil perhitungan nilai koefisien silhouette dapat bervariasi antara -1 hingga 1. Jika nilai koefisien silhouette = 1 berarti objek tersebut sudah berada dalam kelompok yang tepat. Jika nilai koefisien silhouette = 0 maka objek tersebut berada di antara dua kelompok sehingga objek tersebut tidak

jasas harus dimasukkan ke dalam kelompok A atau kelompok B. Jika nilai koefisien silhouette = -1 artinya struktur kelompok yang dihasilkan *overlapping*, sehingga objek tersebut lebih tepat dimasukkan ke dalam kelompok yang lain. Menurut [4], jika nilai *silhouette* yang dihasilkan lebih besar dari 0 dan mendekati 1 maka banyaknya kelompok yang dihasilkan pada proses pengelompokan sudah optimal.

## METODE PENELITIAN

### Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data produksi padi sawah dan padi ladang menurut kabupaten/kota di Provinsi Jawa Tengah yang diambil dari situs resmi Badan Pusat Statistik Jawa Tengah. Pengelompokan kabupaten /kota di Jawa Tengah dengan metode *Density Based Spatial Clustering Algorithm with Noise* (DBSCAN) berdasarkan produksi padi sawah dan padi ladang dilakukan dengan bantuan *software* Matlab.

### Metode Analisis

Analisis yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Menentukan  $Eps$  ( ) dan  $MinPts$
2. Dari data produksi padi sawah dan padi ladang menurut kabupaten/kota di Provinsi Jawa Tengah, diambil secara acak data yang belum dikunjungi
3. Data tersebut kemudian dibaca jumlah tetangganya pada radius .
4. Jika jumlahnya lebih dari atau sama dengan , data akan ditandai sebagai inti dan tetangganya sebagai batas (kecuali yang sudah ditandai sebagai inti), kemudian sebuah kelompok baru akan mulai dibentuk
5. Jika jumlahnya tidak lebih dari atau sama dengan , data tersebut akan ditandai sebagai noise.

6. Tandai data tersebut sebagai data yang sudah dikunjungi.
7. Langkah tersebut dilakukan secara rekursif pada setiap data yang menjadi tetangganya dan belum dikunjungi (pada data produksi padi sawah dan padi ladang menurut kabupaten/kota di Provinsi Jawa Tengah)
8. Dari data produksi padi sawah dan padi ladang menurut kabupaten/kota di Provinsi Jawa Tengah terbentuk beberapa kelompok
9. Melakukan intepretasi dari kelompok-kelompok tersebut.

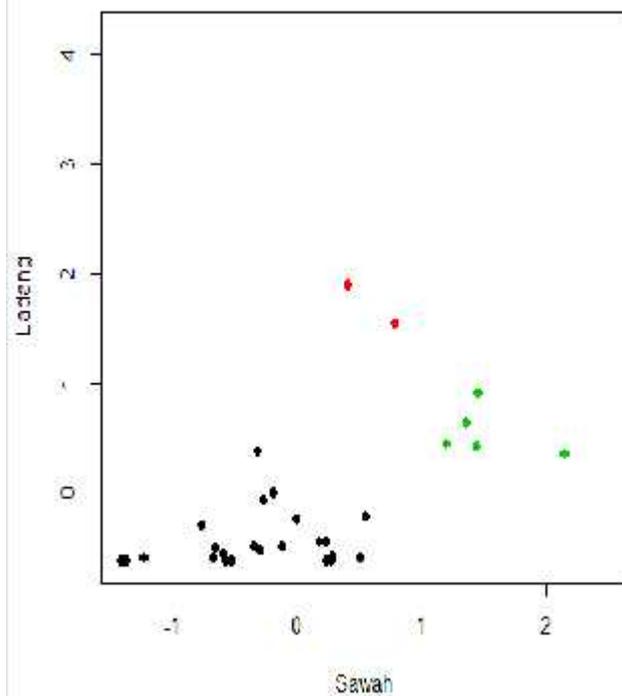
### HASIL PENELITIAN

Pada penelitian ini dilakukan pengelompokan 35 kabupaten /kota di Jawa Tengah menggunakan metode *Density Based Spatial Clustering Algorithm with Noise* (DBSCAN) berdasarkan variabel produksi padi sawah dan padi ladang, dengan bantuan *software* R. Pada penelitian dengan metode DBSCAN ini digunakan minimal points (Minpts) 2 dan epsilon (Eps) 0,9. Sebelum dilakukan pengolahan data, data produksi padi sawah dan padi ladang pada pengelompokan 35 kabupaten /kota di Jawa Tengah distandarkan terlebih dahulu karena mempunyai satuan yang berbeda.

Plot pengelompokan menggunakan DBSCAN dapat dilihat pada Gambar 1. Kelompok 1 ditunjukkan oleh warna hitam yang terdiri dari 26 kabupaten/kota, kelompok 2 ditunjukkan oleh warna merah yang terdiri dari 2 kabupaten/kota, dan kelompok 3 ditunjukkan oleh warna hijau yang terdiri dari 5 kabupaten/kota.

Kelompok 1 terdiri dari Kabupaten Banyumas, Kabupaten Purbalingga, Kabupaten Banjarnegara, Kabupaten

Purworejo, Kabupaten Wonosobo, Kabupaten Magelang, Kabupaten Boyolali, Kabupaten Klaten, Kabupaten Sukoharjo, Kabupaten Karanganyar, Kabupaten Rembang, Kabupaten Kudus, Kabupaten Jepara, Kabupaten Semarang, Kabupaten Temanggung, Kabupaten Kendal, Kabupaten Batang, Kabupaten Pekalongan, Kabupaten Pemalang, Kabupaten Tegal, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, dan Kota Tegal. Kelompok 2 terdiri dari Kabupaten Kebumen dan Kabupaten Blora. Kelompok 3 terdiri dari Kabupaten Sragen, Kabupaten Grobogan, Kabupaten Pati, Kabupaten Demak, dan Kabupaten Brebes. Pada penelitian ini ditemukan 2 *noise*, yaitu Kabupaten Cilacap dan Kabupaten Wonogiri.



**Gambar 1.** Plot pengelompokan 35 kabupaten /kota di Jawa Tengah menggunakan metode (DBSCAN) berdasarkan variabel produksi padi sawah dan padi ladang dengan Eps = 0,9 dan MinPts = 2

**Tabel 1.** Hasil pengelompokan 35 kabupaten /kota di Jawa Tengah menggunakan metode (DBSCAN) berdasarkan variabel produksi padi sawah dan padi ladang dengan Eps = 0,9 dan MinPts = 2

No	Kabupaten/Kota	Padi Sawah	Padi Ladang	Kelompok
1	Kab. Cilacap	248,172	134,682	0
2	Kab. Banyumas	0,22511	-0,44558	1
3	Kab. Purbalingga	-0,29982	-0,52132	1
4	Kab. Banjarnegara	-0,76694	-0,29091	1
5	Kab. Kebumen	0,78190	154,686	2
6	Kab. Purworejo	0,16888	-0,44438	1
7	Kab. Wonosobo	-0,68406	-0,58646	1
8	Kab. Magelang	0,22539	-0,62165	1
9	Kab. Boyolali	-0,19792	0,00145	1
10	Kab. Klaten	0,50182	-0,58911	1
11	Kab. Sukoharjo	0,27230	-0,62165	1
12	Kab. Wonogiri	0,20303	419,435	0
13	Kab. Karanganyar	-0,01158	-0,23941	1
14	Kab. Sragen	134,734	0,63554	3
15	Kab. Grobogan	213,756	0,35735	3
16	Kab. Blora	0,40159	189,512	2
17	Kab. Rembang	-0,32623	0,39056	1
18	Kab. Pati	143,885	0,42393	3
19	Kab. Kudus	-0,66102	-0,50085	1
20	Kab. Jepara	-0,27723	-0,06033	1
21	Kab. Demak	144,271	0,91176	3
22	Kab. Semarang	-0,35429	-0,48713	1
23	Kab. Temanggung	-0,58097	-0,62165	1
24	Kab. Kendal	-0,12034	-0,47813	1
25	Kab. Batang	-0,52970	-0,62165	1
26	Kab. Pekalongan	-0,60118	-0,54617	1
27	Kab. Pemasang	0,54772	-0,20621	1
28	Kab. Tegal	0,27890	-0,57130	1
29	Kab. Brebes	118,858	0,44782	3
30	Kota Magelang	-141,212	-0,62165	1
31	Kota Surakarta	-141,934	-0,61973	1
32	Kota Salatiga	-138,452	-0,62165	1
33	Kota Semarang	-123,258	-0,59214	1
34	Kota Pekalongan	-137,619	-0,62080	1
35	Kota Tegal	-140,740	-0,62165	1

**Tabel 2.** Produksi Padi Sawah dan Padi Ladang untuk Masing – Masing Kelompok

Variabel	Kelompok 1	Kelompok 2	Kelompok 3
Padi Sawah	217546,8	445017,9	647817,1
Padi Ladang	2295,129	31746,29	15949,2

Kelompok 1 adalah kelompok yang mempunyai hasil produksi padi sawah terendah dibandingkan dengan kelompok yang lain. Kelompok 2 adalah kelompok yang mempunyai hasil produksi padi ladang tertinggi dibandingkan dengan kelompok yang lain. Kelompok 3 adalah kelompok yang mempunyai hasil produksi padi sawah tertinggi dibandingkan dengan kelompok yang lain.

## KESIMPULAN

Kelompok 1 terdiri dari Kabupaten/Kota di Provinsi Jawa Tengah selain yang sudah masuk dalam kelompok 2 dan 3, kelompok 1 adalah kelompok yang mempunyai hasil produksi padi sawah terendah dibandingkan dengan kelompok yang lain. Kelompok 2 adalah kelompok yang mempunyai hasil produksi padi ladang tertinggi dibandingkan dengan kelompok yang lain, kelompok 2 terdiri dari Kabupaten Kebumen dan Kabupaten Blora. Kelompok 3 terdiri dari Kabupaten Sragen, Kabupaten Grobogan, Kabupaten Pati, Kabupaten Demak, dan Kabupaten Brebes, adalah kelompok yang mempunyai hasil produksi padi sawah tertinggi dibandingkan dengan kelompok yang lain.

## DAFTAR PUSTAKA

[1] Alfina, T., Santosa, B., Barakbah, A. R., Analisa Perbandingan Metode

Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS), *JURNAL TEKNIK ITS Vol. 1, September 2012, ISSN: 2301-9271*

- [2] Badan Pusat Statistik, 2016, <http://www.bps.go.id>
- [3] Badan Pusat Statistik Jawa Tengah, <http://jateng.bps.go.id>
- [4] Devi, N. M. A. S., Putra, I. K. G. D., Sukarsa, I. M., 2015, Implementasi Metode Clustering DBSCAN pada Proses Pengambilan Keputusan, *Lontar Komputer Vol. 6, No. 3, Desember 2015 ISSN: 2088-154*
- [5] Mumtaz, K., and Duraiswamy, K., An Analysis on Density Based Clustering of Multi Dimensional Spatial Data, *Indian Journal of Computer Science and Engineering, Vol 1 No 1 8-12, ISSN : 0976-5166*
- [6] Prasetyo, E., 2012, *Data Mining - Konsep dan Aplikasi menggunakan MATLAB*, C.V. ANDI OFFSET, Yogyakarta