

SKRIPSI

**PENGELOMPOKAN DATA PENYAKIT GINJAL KRONIS
MENGUNAKAN ALGORITMA DBSCAN**



Oleh :

Richardo Anggara

1822250030

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN REKAYASA
UNIVERSITAS MULTI DATA PALEMBANG
PALEMBANG
2022**

**Fakultas Ilmu Komputer dan Rekayasa
Universitas Multi Data Palembang**

Program Studi Informatika
Skripsi Sarjana Komputer
Semester Genap Tahun 2021/2022

**PENGELOMPOKAN DATA PENYAKIT GINJAL KRONIS
MENGUNAKAN ALGORITMA DBSCAN**

Richardo Anggara

1822250030

Abstrak

Penyakit ginjal kronis adalah penyakit ginjal yang ditandai dengan kerusakan struktural atau fungsional pada ginjal yang berlangsung lebih dari 3 bulan. Penyakit ginjal kronis ditandai dengan satu atau lebih tanda kerusakan ginjal, yaitu albuminuria, sedimen urin abnormal, elektrolit, histologi, struktur ginjal, atau riwayat transplantasi ginjal, dengan penurunan laju filtrasi glomerulus. Pada penelitian ini dilakukan pengelompokan data pasien penyakit ginjal kronis. Tujuan dari penelitian ini adalah mengimplementasikan algoritma *DBSCAN* untuk mengelompokkan data pasien terdiagnosis penyakit ginjal kronis. Dataset yang digunakan adalah *Chronic Kidney Disease* yang berjumlah 400 data medis dengan 24 atribut atau fitur. Metode pengelompokan yang digunakan adalah algoritma *DBSCAN*. Hasil penelitian dievaluasi dengan mencatat banyak jumlah cluster dan noise yang didapatkan. Hasil terbaik ada pada skenario kedua dengan nilai *Silhouette* sebesar 0.15801077304214137 yang mendapatkan *cluster* sebanyak 2 dengan nilai *Epsilon* sebesar 3,5 dan *MinSample* sebanyak 5.

Kata kunci: Penyakit ginjal kronis, *DBSCAN*, *Epsilon*, *Cluster*, *Noise*, *MinSample*, *Silhouette*



BAB 1

PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, rumusan masalah, ruang lingkup, tujuan penelitian, manfaat penelitian, metodologi penelitian dan sistematika penulisan.

1.1 Latar Belakang

Meningkatnya jumlah penderita penyakit ginjal kronis merupakan masalah Kesehatan yang menglobal. Jumlah pengidap penyakit ginjal kronis meningkat karena bertambahnya jumlah penduduk yang berumur lanjut. Berdasarkan hasil dari Global Burden of Disease pada 2010, Penyakit ginjal kronis merupakan pemicu kematian peringkat ke-27 di dunia pada 1990 dan bertambah menjadi urutan ke-16 pada 2020. (Gliselda, 2021)

Penyakit ginjal kronis adalah penyebab utama ke-16 dari tahun kehidupan yang hilang di seluruh dunia. perawatan primer seperti *Screening*, diagnosis, dan manajemen yang tepat oleh dokter diperlukan untuk mencegah Penyakit ginjal kronis yang merugikan, termasuk penyakit kardiovaskular, penyakit ginjal stadium akhir, dan kematian. Penyakit ginjal kronis yang dibiarkan tanpa penanganan yang tepat dapat menyebabkan penyakit ginjal ginjal. Penyakit ginjal kronis harus dicegah dan dikenali dari awal untuk mendapatkan terapi yang efektif. (Gliselda, 2021)

Data mining merupakan prosedur untuk menemukan korelasi, pola, dan tren baru dengan menggali data dengan jumlah besar menggunakan statistik dan teknik matematika yang ada. Banyak investasi dari perusahaan besar telah masuk ke dalam *data mining* namun belum menghasilkan banyak keuntungan. Meskipun data berisi beberapa informasi berharga, keberadaannya masih tersembunyi di gudang data. (Sugiono, Nurdiani, Linawati, Safitri, & Saputra, 2019)

Clustering (Pengelompokan) merupakan proses *clustering* sejumlah data agar terbentuknya beberapa kelas dengan atribut data masing-masing. Salah satu Algoritma *density based clustering* adalah algoritma yang paling efisien untuk mendefinisikan *cluster* data dengan kepadatan yang berbeda *clustering*. (Safitri, Wuryandari, & Rahmawati, 2017)

Density Based Spatial Clustering Algorithm with Noise (DBSCAN) adalah algoritma yang dilandaskan pada kepadatan (*density*) data. DBSCAN merupakan contoh pelopor dalam pengembangan teknik *clustering* berbasis kepadatan atau *clustering* berbasis kepadatan..(Safitri et al., 2017)

DBSCAN tidak perlu mengetahui jumlah kelompok dalam data secara sesukanya seperti pada *K-Means*. Hal ini memberikan keuntungan karena umumnya bentuk dan jumlah kelompok yang sebaiknya diberikan pada data berdimensi tinggi tidak dapat diketahui dengan cara analisis visual data. DBSCAN juga Dapat mengenali *noise* dengan baik. (Nur, 2017)

Berdasarkan penjelasan yang telah diuraikan mengenai penyakit Ginjal Kronis yang harus diberikan perhatian khusus karena menjadi urutan ke 16 penyebab

kematian di dunia menurut Djoko Santoso Kepala badan Kesehatan MUI Jawa Timur. maka peneliti melakukan salah satu metode *Data Mining* yaitu *clustering* pada data pasien terdiagnosis penyakit Ginjal kronis. Penelitian yang dilakukan adalah mengimplementasikan metode *clustering* DBSCAN untuk mengelompokkan data pasien terdiagnosis penyakit ginjal kronis yang memiliki ciri dan karakteristik yang sama dari dataset *Chronic Kidney Disease*.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah didapatkan, maka dapat dirumuskan masalah dari penelitian ini merupakan “Bagaimana mengelompokkan data penyakit ginjal kronis menggunakan algoritma *DBSCAN*”.

1.3 Ruang Lingkup

Ruang lingkup yang digunakan dalam penelitian klasifikasi penyakit ginjal kronis ini yaitu sebagai berikut.

1. Objek yang digunakan berupa data numerik pasien terdiagnosis penyakit ginjal kronis dari dataset *Chronic kidney disease* yang berasal dari <https://www.kaggle.com/mansoordaku/ckdisease>.
2. Metode *clustering* yang digunakan adalah *DBSCAN*.
3. Implementasi kode dilakukan menggunakan *Google Colab* dengan memakai bahasa pemrograman *Python*.
4. *Dataset* yang digunakan berupa *.csv*.
5. Data dinormalisasi menggunakan *min-max*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengimplementasi Algoritma DBScan pada pengelompokan data pasien terdiagnosis penyakit Ginjal kronis.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Memberikan pengetahuan tentang penerapan *Data Mining* pada data pasien terdiagnosis penyakit ginjal kronis.
2. Memberikan pengetahuan tentang penggunaan Algoritma DBSCAN.
3. Hasil dari penelitian diharapkan dapat menjadi acuan bagi dokter untuk memberikan *treatment* atau penanganan yang sama terhadap pasien terdiagnosis penyakit ginjal kronis yang sesuai dengan kelompok *cluster* dengan ciri-ciri atau karakteristik yang sama.

1.6 Sistematika Penulisan

Sistematika penulisan laporan skripsi yang akan dilakukan antara lain sebagai berikut.

BAB 1 PENDAHULUAN

Bab ini adalah penjelasan mengenai latar belakang, rumusan masalah, ruang lingkup, tujuan penelitian, manfaat, metodologi penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini adalah penjelasan mengenai teori implementasi metode *DBSCAN* dalam pengelompokan data pasien terdiagnosis penyakit ginjal kronis.

BAB 3 RANCANGAN PENELITIAN

Bab ini adalah penjelasan mengenai kebutuhan perangkat keras dan lunak, serta perancangan metode penelitian yang digunakan.

BAB 4 HASIL DAN BAHASAN

Bab ini adalah penjelasan mengenai analisis dan pembahasan hasil implementasi metode *DBSCAN* pada pengelompokan data pasien terdiagnosis penyakit ginjal kronis yang sudah dibuat.

BAB 5 KESIMPULAN DAN SARAN

Bab ini adalah penjelasan mengenai kesimpulan penelitian yang telah dilakukan dan saran yang berguna bagi pengembangan penelitian selanjutnya.



DAFTAR PUSTAKA

- Aisara, S., Azmi, S., & Yanni, M. (2018). Gambaran Klinis Penderita Penyakit Ginjal Kronik yang Menjalani Hemodialisis di RSUP Dr. M. Djamil Padang. *Jurnal Kesehatan Andalas*, 7(1), 42. <https://doi.org/10.25077/jka.v7i1.778>
- Aldhyani, T. H. H., Alshebami, A. S., & Alzahrani, M. Y. (2020). Soft Clustering for Enhancing the Diagnosis of Chronic Diseases over Machine Learning Algorithms. *Journal of Healthcare Engineering*, 2020. <https://doi.org/10.1155/2020/4984967>
- Ashari, B. S., Otniel, S. C., & Rianto. (2019). Perbandingan Kinerja K-Means Dengan DSCAN Untuk Metode Clustering Data Penjualan Online Retail. *Jurnal Siliwangi*, 5(2), 72–77. Retrieved from <http://jurnal.unsil.ac.id/index.php/jssainstek/article/view/1283>
- Chen, Y., Zhou, L., Bouguila, N., Wang, C., Chen, Y., & Du, J. (2021). BLOCK-DBSCAN: Fast clustering for large scale data. *Pattern Recognition*, 109. <https://doi.org/10.1016/j.patcog.2020.107624>
- Gliselda, V. K. (2021). Diagnosis dan Manajemen Penyakit Ginjal Kronis (PGK). *Jurnal Medika Utama*, 2(04 Juli), 1135–1141.
- Goleman, daniel; boyatzis, Richard; Mckee, A. (2019). Algoritma DBSCAN dan Contoh Perhitungannya. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
- Gupta, P. (2021). *Practical Data Science with Jupyter: Explore Data Cleaning, Pre-processing, Data Wrangling, Feature Engineering and Machine Learning using Python and Jupyter*.
- Hahsler, M., Piekenbrock, M., & Doran, D. (2019). Dbscan: Fast density-based clustering with R. *Journal of Statistical Software*, 91(1).

<https://doi.org/10.18637/jss.v091.i01>

Id, I. (2017). Modifikasi DBSCAN (Density-Based Spatial Clustering with Noise) Pada Objek 3 Dimensi. *Jurnal Komputer Terapan*, 3(1), 41–52.

Kuntjoro, D. A., Setiawan, B. D., & Perdana, R. S. (2018). Algoritme Genetika Untuk Optimasi K-Means Clustering Dalam Pengelompokan Data Tsunami. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(10), 3865–3872. Retrieved from <http://j-ptiik.ub.ac.id>

Mahboob, T., Ijaz, A., Shahzad, A., & Kalsoom, M. (2019). Handling Missing Values in Chronic Kidney Disease Datasets Using KNN, K-Means and K-Medoids Algorithms. *ICOSST 2018 - 2018 International Conference on Open Source Systems and Technologies, Proceedings*, 76–81. <https://doi.org/10.1109/ICOSST.2018.8632179>

Oluyombo, R., Olamoyegun, M. A., Ayodele, O. E., Akinwusi, P. O., & Akinsola, A. (2017). Clustering of chronic kidney disease and cardiovascular risk factors in South-West Nigeria. *Journal of Nephropathology*, 6(3), 196–203. <https://doi.org/10.15171/jnp.2017.33>

Ordila, R., Wahyuni, R., Irawan, Y., & Yulia Sari, M. (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. <https://doi.org/10.33060/jik/2020/vol9.iss2.181>

Putri, M. M., Dewi, C., Permata Siam, E., Asri Wijayanti, G., Aulia, N., & Nooraeni, R. (2021). Komparasi DBSCAN dan K-Means Clustering pada Pengelompokan Status Desa di Jawa Tengah Tahun 2020, 17(3), 394–404. <https://doi.org/10.20956/j.v17i3.11704>

Qadrini, L. (2020). Metode K-Means dan DBSCAN pada Pengelompokan Data Dasar Kompetensi Laboratorium ITS Tahun 2017. *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika*, 13(2), 5–11. <https://doi.org/10.36456/jstat.vol13.no2.a2886>

- Risman, Syaripuddin, & Suyitno. (2019). Implementasi Metode DbSCAN Pada Pengelompokan Kabupaten/Kota Di Pulau Kalimantan Berdasarkan Indikator Kesejahteraan Rakyat. *Prosiding Seminar Nasional Matematika*, 22–28.
- Safitri, D., Wuryandari, T., & Rahmawati, R. (2017). Metode DbSCAN Untuk Pengelompokan Kabupaten / Kota Di Provinsi Jawa Tengah. *Statistika*, 5(1), 8–13.
- Sari, B. N., & Primajaya, A. (2019). Penerapan Clustering DbSCAN Untuk Pertanian Padi Di Kabupaten Karawang. *Jurnal Informatika Dan Komputer*, 4(1), 28–34. Retrieved from www.mapcoordinates.net/en.
- Sugiono, Nurdiani, S., Linawati, S., Safitri, R. A., & Saputra, E. P. (2019). Pengelompokan Perilaku Mahasiswa Pada Perkuliahan E-Learning dengan K-Means Clustering. *Jurnal Kajian Ilmiah*, 19(2), 126–133.
- Suntoro, J. (2019). 22-DATA MINING Algoritma dan Implementasi Menggunakan Bahasa Pemrograman PHP. *DATA MINING Algoritma Dan Implementasi Menggunakan Bahasa Pemrograman PHP*, 9(9), 259–278.