

Entry Kartu Bimbingan Tugas Akhir



NIM : 2112501560

Nama Mahasiswa : M. Bintang Akram

Dosen Pembimbing : Yudi Santoso

Judul Skripsi : PENERAPAN METODE K-MEANS CLUSTERING UNTUK PENGELOMPOKAN RESIKO PASIEN PENYAKIT GINJAL KRONIK

Hari / Tanggal / Jam : 27/08/2025 09:24

Isikan Materi :

SIMPAN

Kartu Bimbingan Mahasiswa

No	Tanggal	Materi	Action
1	21-APR-2025 13:15:00	konsultasi terkait dataset yang ingin digunakan	EDIT DELETE
2	29-APR-2025 11:00:00	konsultasi terkait judul dan dataset sudah di setuju	EDIT DELETE
3	03-MAY-2025 13:15:00	konsultasi tentang latar belakang dan bab 1	EDIT DELETE
4	14-MAY-2025 14:15:00	konsultasi dan pemeriksaan menyeluruh tentang bab 1	EDIT DELETE
5	20-MAY-2025 13:15:00	konsultasi tentang bab 2 (noted : perbanyak informasi yang sesuai dengan topik yang dipilih dan perbanyak refrensi jurnal)	EDIT DELETE
6	10-JUN-2025 08:22:00	konsultasi tentang bab 3 (perbandingan dengan penelitian terdahulu)	EDIT DELETE
7	20-JUN-2025 13:15:00	konsultasi mengenai bab 4 (penulisan table, input gambar)	EDIT DELETE
8	02-JUL-2025 08:15:00	konsultasi bab 4 dan 5 mengenai penulisan tabel, input gambar, isi berita acara dan konsultasi mengenai sidang TA	EDIT DELETE



**UNIVERSITAS
BUDI LUHUR**

Kampus Pusat : Jl. Raya Ciledug - Petukangan Utara - Jakarta Selatan 12260
Telp : 021-5853753 (hunting), Fax : 021-5853489, <http://www.budiluhur.ac.id>

FAKULTAS TEKNOLOGI INFORMASI
FAKULTAS EKONOMI DAN BISNIS
FAKULTAS ILMU SOSIAL DAN STUDI GLOBAL
FAKULTAS TEKNIK
FAKULTAS KOMUNIKASI DAN DESAIN KREATIF

BERITA ACARA SIDANG PENDADARAN TUGAS AKHIR

S/UBL/FTI/1323/VII/25

Pada hari ini, Rabu 16 Juli 2025 telah dilaksanakan Ujian Sidang Pendadaran Tugas Akhir sebagai berikut:

Judul : PENERAPAN METODE K-MEANS CLUSTERING UNTUK PENGELOMPOKAN RESIKO PASIEN
PENYAKIT GINJAL KRONIK

Nama : M. Bintang Akram

NIM : 2112501560

Dosen Pembimbing : Yudi Santoso, S.Kom., M.Kom.

Berdasarkan penilaian pada Presentasi, Penulisan, Penguasaan Materi, Penguasaan Program / Prototipe, Kompleksitas maka Mahasiswa tersebut di atas dinyatakan:

LULUS

dengan nilai angka: 92 huruf: A

Mahasiswa tersebut di atas wajib menyerahkan hasil perbaikan tulisan Tugas Akhir dalam bentuk terjilid sesuai dengan Panduan Perbaikan Tugas Akhir, selambat-lambatnya Rabu 30 Juli 2025.

Panitia Penguji:

1. Ketua Bruri Trya Sartana, S.Kom., M.M., M.Kom.
2. Anggota Ir. Bullion Dragon Andah, M.Sc, M.Kom.
3. Moderator Yudi Santoso, S.Kom., M.Kom.

Keterangan:

Nilai Huruf: A:85-100 A-:80-84,99 B+:75-79,99 B:70-74,99 B-:65-69,99 C:60-64,99 D:40-59,99 E-:0-39,99

**PENERAPAN METODE K-MEANS CLUSTERING UNTUK
PENGELOMPOKAN RISIKO PASIEN PENYAKIT GINJAL
KRONIK**

TUGAS AKHIR



Oleh:

**M BINTANG AKRAM
NIM: 2112501560**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2025**



LEMBAR PENGESAHAN

Nama	: M. Bintang Akram
Nomor Induk Mahasiswa	: 2112501560
Program Studi	: Sistem Informasi
Bidang Peminatan	: Data Science
Jenjang Studi	: Strata 1
Judul	: PENERAPAN METODE K-MEANS CLUSTERING UNTUK PENGELOMPOKAN RESIKO PASIEN PENYAKIT GINJAL KRONIK



Laporan Tugas Akhir ini telah disetujui, disahkan dan direkam secara elektronik sehingga tidak memerlukan tanda tangan tim penguji.

Jakarta, Rabu 16 Juli 2025

Tim Penguji:

Ketua	: Bruri Trya Sartana, S.Kom., M.M., M.Kom
Anggota	: Ir. Bullion Dragon Andah, M.Sc, M.Kom
Pembimbing	: Yudi Santoso, S.Kom., M.Kom.
Ketua Program Studi	: Bima Cahya Putra, S.Kom., M.Kom.

ABSTRAK

PENERAPAN METODE K-MEANS CLUSTERING UNTUK PENGELOMPOKAN RISIKO PASIEN PENYAKIT GINJAL KRONIK

Oleh: M Bintang Akram (2112501560)

Penyakit ginjal kronik merupakan salah satu penyakit tidak menular dengan tingkat morbiditas tinggi yang memerlukan penanganan berbasis data untuk menentukan prioritas pasien secara tepat sasaran. Permasalahan yang diangkat dalam penelitian ini adalah bagaimana mengelompokkan pasien penyakit ginjal kronik ke dalam kelompok risiko berdasarkan atribut klinis yang dimiliki. Tujuan dari penelitian ini adalah untuk menerapkan metode K-Means Clustering sebagai pendekatan segmentasi risiko pasien, serta membandingkan performanya dengan algoritma *X-Means*. Metode penelitian mengikuti kerangka kerja CRISP-DM, yang mencakup pemahaman masalah, pengumpulan dan pembersihan data, transformasi atribut kategorikal menjadi numerik menggunakan Label Encoding, normalisasi data dengan metode Min-Max, pemodelan menggunakan algoritma *K-Means* dan *X-Means*, evaluasi hasil klasterisasi, serta pembahasan hasil. Penentuan jumlah klaster optimal pada K-Means dilakukan menggunakan metode Elbow, yang menunjukkan hasil terbaik pada $k=2$, sedangkan *X-Means* secara otomatis menentukan jumlah klaster optimal. Hasil klasterisasi dari kedua algoritma menunjukkan pembentukan dua kelompok pasien, yaitu pasien dengan kondisi klinis ringan hingga sedang (Cluster 0) dan pasien dengan kondisi klinis berat (Cluster 1). Evaluasi menggunakan Davies-Bouldin Index (DBI) menunjukkan bahwa kedua algoritma menghasilkan nilai DBI yang sama, yaitu 1.119, pada jumlah dua klaster. Meskipun demikian, *K-Means* dipilih sebagai model utama dalam penelitian ini karena memiliki keunggulan dari segi efisiensi komputasi, kemudahan implementasi, serta kesesuaian dengan tujuan penelitian yang menekankan eksplorasi eksplisit terhadap jumlah klaster. Kesimpulan dari penelitian ini adalah bahwa metode *K-Means* efektif dalam mengelompokkan pasien penyakit ginjal kronik ke dalam dua kelompok risiko yang berbeda, yang dapat digunakan sebagai dasar penentuan prioritas penanganan pasien secara lebih tepat sasaran dan berbasis data.

Kata Kunci : *K-Means Clustering*, penyakit ginjal kronik, Davies-Bouldin Index, segmentasi risiko pasien

DAFTAR ISI

LEMBAR PENGESAHAN	i
ABSTRAK.....	ii
SURAT PERNYATAAN TIDAK PLAGIAT DAN PERSETUJUAN PUBLIKASI.....	iii
KATA PENGANTAR	iv
DAFTAR TABEL	vi
DAFTAR GAMBAR	vii
DAFTAR ISI.....	viii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Masalah.....	2
1.2.1 Identifikasi Masalah.....	2
1.2.2 Batasan Masalah	3
1.2.3 Rumusan Masalah.....	3
1.3 Tujuan dan Manfaat Penulisan.....	4
1.3.1 Tujuan	4
1.3.2 Manfaat	4
1.4 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Teori-Teori Umum	6
2.1.1 Pengertian <i>Data Mining</i>	6
2.1.2 Proses Penambangan Data (<i>Data Mining</i>).....	6
2.1.3 <i>K-Means Clustering</i>	7
2.1.4 Metode <i>Elbow</i>	7
2.2 Teori Khusus	7
2.2.1 Klasterisasi.....	7
2.2.2 Pengertian Ginjal	8
2.2.3 Penyakit Ginjal Kronik	8
2.2.4 Faktor Resiko Penyakit Ginjal Kronik.....	9
2.3 <i>Tools</i> Yang Digunakan	9
2.3.1 Google Collabs	9
2.3.2 Rapid Miner	10
2.3.3 Microsoft Excel.....	10
2.4 Studi Literatur	10
BAB III METODOLOGI PENELITIAN	12
3.1 Tahapan Penelitian	12
3.1.1 Studi Literatur	13
3.1.2 Identifikasi Masalah.....	13
3.1.3 Analisa Masalah.....	13
3.1.4 Kesimpulan Dan Saran	14
3.2 Perbedaan dari penelitian sebelumnya.....	14
3.3 <i>Data Preprocessing</i>	15
3.4 Penentuan Data Latih dan Data Uji	16
3.6 Teknik Pengujian	17
BAB IV ANALISIS DAN PEMBAHASAN	18

4.1	Pengumpulan Data	18
4.2	Data <i>Preprocessing</i> / PraPemosesan Data	19
4.2.1	Memahami Situasi dan Menetapkan Tujuan yang ingin Dicapai ..	20
4.2.2	<i>Data Selection</i>	20
4.2.3	<i>Data Cleaning</i>	23
4.2.4	<i>Data Integration</i>	27
4.2.5	Transformasi Data	27
4.2.6	Normalisasi Data	29
4.3	Pemodelan	32
4.3.1	Hasil Komparasi Model	32
4.3.2	Penyajian Model Terbaik	33
4.3.3	Perhitungan Algoritma Terpilih	40
4.4	Pengujian	46
BAB V KESIMPULAN		47
5.1	Kesimpulan	47
5.2	Saran	47
DAFTAR PUSTAKA		49
LAMPIRAN A DATASET AWAL		51
LAMPIRAN B DATASET HASIL KLASTERISASI		53
LAMPIRAN C		55

DAFTAR PUSTAKA

- 3th ANNUAL REPORT OF INDONESIAN RENAL REGISTRY 2020. (n.d.).
www.indonesianrenalregistry.org
- Alwiyah, F., Rudyanto, W., Indria Anggraini, D., & Windarti, I. (2024). Anatomi dan Fisiologi Ginjal: Tinjauan Pustaka. In *Tinjauan Pustaka Medula* / (Vol. 14).
- Amri, Z., Muhammad Rodi, M. Nurul Wathani, Amir Bagja, & Zulkipli. (2025). Prediksi Diabetes Menggunakan Algoritma K-Nearest (KNN) Teknik SMOTE-ENN. *Infotek: Jurnal Informatika Dan Teknologi*, 8(1), 193–204.
<https://doi.org/10.29408/jit.v8i1.27975>
- Aulia Shabrina, S., Saftarina, F., & Anggileo Pramesona, B. (n.d.). *Faktor Risiko Penyakit Ginjal Kronik Pada Pasien Diabetes*.
- Debal, D. A., & Sitote, T. M. (2022). Chronic kidney disease prediction using machine learning techniques. *Journal of Big Data*, 9(1). <https://doi.org/10.1186/s40537-022-00657-5>
- Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi K-Means Clustering untuk Pengelompokan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi Dan Informasi*, 12(1), 64–77. <https://doi.org/10.34010/jati.v12i1.6674>
- Falah, M. H. , Setyawati, T. , Walanda, R. M. , & Putrie, I. R. (n.d.). *HUBUNGAN KADAR KREATININ DENGAN HEMOGLOBIN PADA PASIEN GAGAL GINJAL KRONIK (GGK) DI RSUD UNDATA PROVINSI SULAWESI TENGAH TAHUN 2022*.
- Febyolla, C. L., Pardilawati, C. Y., Junando, M., & Damayanti, E. (2025). Article Review: Faktor Risiko Terjadinya Gagal Ginjal Kronik Di Indonesia. *Jurnal Farmasi SYIFA*, 3(1), 50–57. <https://doi.org/10.63004/jfs.v3i1.646>
- Janani J, & Sathyaraj R. (2021). Diagnosing Chronic Kidney Disease Using Hybrid Machine Learning Techniques. In *Turkish Journal of Computer and Mathematics Education* (Vol. 12, Issue 13).
<https://doi.org/https://doi.org/10.17762/turcomat.v12i13.9943>
- Jhee, J. H., Hwang, S. D., Song, J. H., & Lee, S. W. (2018). Upper normal serum creatinine concentrations as a predictor for chronic kidney disease: Analysis of 14 years' korean genome and epidemiology study (KoGES). *Journal of Clinical Medicine*, 7(11). <https://doi.org/10.3390/jcm7110463>
- Krisnanta Wijaya, A., Syifa, R., Rahmadianto, I. N., & Hapsari, R. K. (2024). *Seminar Nasional Informatika Bela Negara (SANTIKA) Identifikasi Penyakit Ginjal Kronis Menggunakan Algoritma K-Nearest Neighbour (k-NN)*.
<https://www.kaggle.com/>
- Li, Z., He, R., Wang, Y., Qu, Z., Liu, J., Yu, R., & Yang, S. (2025). Global trends of chronic kidney disease from 1990 to 2021: a systematic analysis for the global burden of

- disease study 2021. *BMC Nephrology*, 26(1), 385.
<https://doi.org/10.1186/s12882-025-04309-7>
- Nazar, Ridwan. (2024). *Implementasi Pemrograman Python Menggunakan Google Colab*.
- Rizky, I. I. M., Irianto, S. Y., & Sriyanto. (2023). *Perbandingan Kinerja Algoritma Naive Bayes, Support Vector Machine dan Random forest untuk Prediksi Penyakit Ginjal Kronis*. <https://jurnal.darmajaya.ac.id/index.php/PSND/article/view/3832>
- Rohendi, G. F., Suarna, N., & Lestari, G. D. (2023). Analisis Rumah Tidak Layak Huni Menggunakan Algoritma X-Means. *Jurnal Janitra Informatika Dan Sistem Informasi*, 3(1), 18–29. <https://doi.org/10.25008/janitra.v3i1.151>
- Syahputra, E., Laoli, E. K., Alyah, J., HSB, E. Y. B., Estra, E. Y., br. Tumorang, & Nababan, T. (2022). *Dukungan Keluarga Berhubungan Dengan Kualitas Hidup Pasien Gagal Ginjal Kronik Yang Menjalani Terapi Hemodialisa*.
<http://jurnal.globalhealthsciencegroup.com/index.php/JPPP>
- Vania, P., & Nurina Sari, B. (2023). Perbandingan Metode Elbow dan Silhouette untuk Penentuan Jumlah Kluster yang Optimal pada Clustering Produksi Padi menggunakan Algoritma K-Means. *Jurnal Ilmiah Wahana Pendidikan*, 9(21), 547–558. <https://doi.org/10.5281/zenodo.10081332>
- Wardana, M. A., & Suherman, S. (2025). Penerapan Data Mining Pengelompokan Data Penjualan Motor di PT. Nusantara Surya Sakti Soppeng Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika (JISTI)*, 8(1), 123–132. <https://doi.org/10.57093/jisti.v8i1.282>
- Yudha Crisanto, E., Djamaludin, D., Yulendasari, R., Purnama Sari, R., Triyono, & Umsani. (2022). Penyuluhan kesehatan tentang perilaku sehat pasien gagal ginjal kronik (GGK). In *JOURNAL OF Public Health Concerns* (Vol. 2, Issue 2). <https://e-jurnal.ipphorr.com/index.php/phc>
- Yusri, R., Edriati, S., Yuhendri, R., & PGRI Sumatera Barat, S. (2020). *Pelatihan Microsoft Office Excel Sebagai Upaya Peningkatan Kemampuan Mahasiswa Dalam Mengolah Data*. 2(1), 32–37.
<https://doi.org/10.22202/JR.2020.V1i2.3914>