



UNIVERSITAS BUDI LUHUR

FAKULTAS TEKNOLOGI INFORMASI

Kartu Bimbingan Tugas Akhir

NIM: 2211501743

Nama: Galang Danang Pramudya

Pembimbing: Dolly Virgiana Shaka Yudha Sakti, S.Kom., M.Kom.

No.	Tanggal	Materi
1	21-04-2026	Penentuan topik dan judul
2	22-05-2026	Pembahasan mengenai data
3	29-05-2026	Membahas Bab 1
4	03-06-2026	Membahas Bab 2
5	08-06-2026	pembahasan mengenai data dan indikator perhitungan
6	18-06-2026	Pembahasan mengenai tata letak dan user interface website
7	22-06-2026	Membahan BAB 3 LRS, Flow chart
8	29-06-2026	Demo web



BERITA ACARA SIDANG PENDADARAN TUGAS AKHIR

S/UBL/FTI/0597/VII/26

Pada hari ini, Rabu 15 Juli 2026 telah dilaksanakan Ujian Sidang Pendadaran Tugas Akhir sebagai berikut:

Judul: IMPLEMENTASI DATA MINING UNTUK KLASTERISASI TINGKAT KERAWANAN BANJIR DI WILAYAH DKI JAKARTA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING BERBASIS WEB

Nama : Galang Danang Pramudya
NIM : 2211501743
Dosen Pembimbing : Dolly Virgiani Shaka Yudha Sakti, S.Kom., M.Kom.

Berdasarkan penilaian pada Presentasi + Demo, Penulisan, Penguasaan Materi, Penguasaan Program maka Mahasiswa tersebut di atas dinyatakan:

LULUS

dengan nilai angka : **83** 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 29 Juli 2026.

Panitia Penguji:

- | | |
|-------------|--|
| 1 Ketua | Wahyu Pramusinto, S.Kom., M.Kom. |
| 2 Anggota | Dr. Indra, S.Kom., M.T.I. |
| 3 Moderator | Dolly Virgiani Shaka Yudha Sakti, 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

**IMPLEMENTASI *DATA MINING* UNTUK KLASTERISASI
TINGKAT KERAWANAN BANJIR DI WILAYAH DKI
JAKARTA MENGGUNAKAN ALGORITMA K-MEANS
CLUSTERING BERBASIS WEB**

TUGAS AKHIR



Oleh:

**GALANG DANANG PRAMUDYA
NIM : 2211501743**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2026**

**IMPLEMENTASI *DATA MINING* UNTUK KLASTERISASI
TINGKAT KERAWANAN BANJIR DI WILAYAH DKI
JAKARTA MENGGUNAKAN ALGORITMA K-MEANS
CLUSTERING BERBASIS WEB**

TUGAS AKHIR



Oleh:

**GALANG DANANG PRAMUDYA
NIM : 2211501743**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2026**

**IMPLEMENTASI *DATA MINING* UNTUK KLASTERISASI
TINGKAT KERAWANAN BANJIR DI WILAYAH DKI
JAKARTA MENGGUNAKAN ALGORITMA K-MEANS
CLUSTERING BERBASIS WEB**

**Diajukan Untuk Memenuhi Satu Persyaratan Memperoleh Gelar
Sarjana Komputer (S.KOM)**

TUGAS AKHIR



Oleh:

**GALANG DANANG PRAMUDYA
NIM : 2211501743**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2026**



LEMBAR PENGESAHAN

Nama	: Galang Danang Pramudya
Nomor Induk Mahasiswa	: 2211501743
Program Studi	: Teknik Informatika
Bidang Peminatan	: Programming Expert
Jenjang Studi	: Strata 1
Judul	: IMPLEMENTASI DATA MINING UNTUK KLASTERISASI TINGKAT KERAWANAN BANJIR DI WILAYAH DKI JAKARTA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING BERBASIS WEB



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

Jakarta, Rabu 15 Juli 2026

Tim Penguji:

Ketua	: Wahyu Pramusinto, S.Kom., M.Kom.
Anggota	: Dr. Indra, S.Kom., M.T.I.
Pembimbing	: Dolly Virgiani Shaka Yudha Sakti, S.Kom., M.Kom.
Ketua Program Studi	: Dr. Indra, S.Kom., M.T.I.

ABSTRAK

IMPLEMENTASI DATA MINING UNTUK KLASTERISASI TINGKAT KERAWANAN BANJIR DI WILAYAH DKI JAKARTA MENGGUNAKAN ALGORITMA K-MEANS *CLUSTERING* BERBASIS WEB

Oleh : Galang Danang Pramudya (2211501743)

Banjir merupakan bencana alam yang paling sering terjadi di DKI Jakarta, menimbulkan kerugian material, terganggunya aktivitas warga, dan bertambahnya korban terdampak setiap tahun. Penelitian ini menerapkan algoritma K-Means *clustering* untuk mengelompokkan tingkat kerawanan banjir di DKI Jakarta berdasarkan data historis BPBD DKI Jakarta periode 2021–2025, sekaligus membangun *platform* berbasis web untuk menyajikan hasil klasterisasi secara interaktif. Pengolahan data mengikuti kerangka *Knowledge Discovery in Databases* (KDD), meliputi seleksi data, *preprocessing*, transformasi, *data mining*, dan evaluasi, dengan lima indikator jumlah kejadian banjir, jumlah kepala keluarga terdampak, jumlah jiwa terdampak, jumlah pengungsi, dan rata-rata ketinggian air. Setelah dinormalisasi menggunakan *Min-Max Normalization*, data diproses dengan K-Means pada jumlah klaster optimal $K=4$, menghasilkan kategori Rendah, Sedang, Tinggi, dan Sangat Tinggi. Evaluasi dilakukan dengan membandingkan *Euclidean Distance* dan *Manhattan Distance* menggunakan *Silhouette Score* dan *Purity Score*. Hasil pengujian menunjukkan *Silhouette Score* sebesar 0,4578 (*Euclidean*) dan 0,5618 (*Manhattan*), sedangkan *Purity Score* sebesar 69,61% (*Euclidean*) dan 51,32% (*Manhattan*). Berdasarkan *Purity Score* yang lebih selaras dengan kategori kerawanan resmi BPBD, *Euclidean Distance* dipilih sebagai metode jarak terbaik dalam sistem ini.

Kata kunci: *Banjir, K-Means clustering, Data Mining, KDD, Mitigasi Bencana, DKI Jakarta, Web.*

DAFTAR ISI

LEMBAR PENGESAHAN.....	i
ABSTRAK.....	ii
SURAT TIDAK PLAGIAT.....	iii
SURAT PERNYATAAN PENGGUNAAN AI GENERATIF.....	iv
KATA PENGANTAR.....	v
DAFTAR TABEL.....	vi
DAFTAR GAMBAR.....	viii
DAFTAR SIMBOL.....	ix
DAFTAR ALGORITMA.....	xi
DAFTAR ISI.....	xii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Perumusan Masalah.....	1
1.3. Batasan Masalah.....	2
1.4. Tujuan.....	2
1.5. Manfaat.....	3
1.6. Sistematika Penulisan.....	3
BAB II LANDASAN TEORI.....	5
2.1. Banjir.....	5
2.2. <i>Data Mining</i>	5
2.3. <i>Knowledge Discovery in Database (KDD)</i>	6
2.4. Normalisasi Data.....	7
2.5. Algoritma K-Means.....	8
2.6. Klasterisasi.....	8
2.7. Evaluasi Kualitas kluster.....	9
2.8. Studi Literatur.....	11
2.9. Perbandingan dan Keterkaitan Antara Jurnal.....	16
BAB III METODOLOGI PENELITIAN.....	18
3.1. Data Penelitian.....	18
3.2. Metode Pembandingan.....	20
3.3. Penerapan Metode.....	20
3.3.1. Tahap Pengumpulan Data.....	21
3.3.2. Seleksi Data.....	22
3.3.3. <i>Preprocessing</i> Data.....	23
3.3.4. <i>Data Mining</i>	24
3.3.5. Evaluasi Hasil <i>clustering</i>	25
3.4. Rancangan Pengujian.....	26
3.4.1. Normalisasi Min Max.....	26
3.4.2. <i>Euclidean Distance</i>	26
3.4.3. <i>Manhattan Distance</i>	26
3.4.4. Pengelompokan Data.....	27
3.4.5. Pusat kluster Baru.....	27
3.4.6. <i>Elbow Method</i>	27
3.4.7. <i>Purity</i>	28
3.4.8. <i>Silhouette Score</i>	28

3.5. Rancangan Basis Data.....	28
3.5.1. <i>Logical Record Structured (LRS)</i>	29
3.5.2. Spesifikasi Basis Data.....	30
3.6. Rancangan Menu.....	42
3.7. Rancangan Layar.....	42
3.7.1. Halaman Publik.....	42
3.7.2. <i>Login</i>	43
3.7.3. <i>Dashboard Admin</i>	44
3.7.4. <i>Import Data</i>	44
3.7.5. <i>Preprocessing Data</i>	45
3.7.6. <i>Elbow Method</i>	47
3.7.7. Proses K-means.....	48
3.7.8. Perbandingan Jarak.....	48
3.7.9. Pelabelan klaster.....	49
3.7.10. Hasil klaster.....	50
3.7.11. Laporan.....	50
BAB IV HASIL DAN PEMBAHASAN.....	52
4.1. Lingkungan Percobaan.....	52
4.1.1. Spesifikasi <i>Hardware</i>	52
4.1.2. Spesifikasi <i>Software</i>	52
4.1.3. <i>Deployment Diagram</i>	52
4.2. Implementasi Metode.....	53
4.2.1. Data.....	53
4.2.3. Penentuan Jumlah klaster.....	56
4.2.4. Penentuan Centroid Awal.....	56
4.2.5. Perhitungan Jarak Data.....	57
4.2.6. Pengelompokan Data.....	61
4.2.7. Perhitungan Centroid Baru.....	62
4.2.8. Perhitungan Purity.....	65
4.3. <i>Flowchart</i>	7
4.3.1. <i>Flowchart</i> Keseluruhan Sistem.....	79
4.3.2. <i>Flowchart</i> Publik.....	79
4.3.3. <i>Flowchart Login</i>	80
4.3.4. <i>Flowchart Import Data</i>	81
4.3.5. <i>Flowchart Preprocessing Data</i>	82
4.3.6. <i>Flowchart Elbow Method</i>	84
4.3.7. <i>Flowchart</i> Proses K-means.....	84
4.3.8. <i>Flowchart</i> Perbandingan Jarak.....	85
4.3.9. <i>Flowchart</i> Labeling klaster.....	86
4.3.10. <i>Flowchart</i> Hasil klaster.....	87
4.3.11. <i>Flowchart</i> Laporan.....	88
4.4. Algoritma.....	88
4.4.1. Algoritma Keseluruhan Sistem.....	89
4.4.2. Algoritma Publik.....	90
4.4.3. Algoritma <i>Preprocessing Data</i>	91
4.4.4. Algoritma Proses K-means <i>clustering</i>	92

4.4.5. Algoritma Hasil kluster.....	93
4.5. Tampilan layar.....	94
4.5.1. Halaman Peta Publik.....	94
4.5.2. Halaman <i>Login</i>	95
4.5.3. <i>Dashboard</i> Admin.....	95
4.5.4. <i>Import</i> Data.....	96
4.5.5. <i>Preprocessing</i> Data.....	96
4.5.6. <i>Elbow Method</i>	98
4.5.7. Proses K-means.....	98
4.5.8. Perbandingan Jarak.....	99
4.5.9. Pelabelan kluster.....	99
4.5.10. Hasil kluster.....	100
4.5.11. Laporan.....	100
BAB V PENUTUP.....	102
5.1. Kesimpulan.....	102
5.2. Saran.....	103
DAFTAR PUSTAKA.....	104
LAMPIRAN.....	107

DAFTAR PUSTAKA

- Amalina, T., Bima, D., Pramana, A., & Sari, B. N. (2022). Metode K-Means clustering Dalam Pengelompokan Penjualan Produk Frozen Food. *Jurnal Ilmiah Wahana Pendidikan*, 8(15), 574–583. <https://doi.org/10.5281/zenodo.7052276>
- Anzari, Y., Puriza, Y., Akbar, N., Abdillah, N., Dahwanu, O., & Syawal, E. A. (2025). Penerapan algoritma K-Means untuk klasterisasi pola iklim studi kasus: Provinsi Jambi periode 2020-2024. *Djtechno: Jurnal Teknologi Informasi*, 6(3), 1037–1052.
- Artana, P. N. P., Mandyartha, E. P., & S, M. H. P. (2023). Penerapan data mining pada algoritma hierarchical clustering tentang pengelolaan mitra perjalanan wisatawan Bali backpacker. *JATI: Jurnal Mahasiswa Teknik Informatika*, 7(4), 2903–2909.
- Arvi, M. D., Sibarani, R. S. Y., Tanjung, Y. I., & Fairuz, T. (2025). Analisis faktor penyebab bencana banjir di kota-kota besar Indonesia: Studi kasus analisis banjir berbasis literasi. *Indonesian Journal of Emerging Trends in Community Engagement*, 3(1), 1–8. <https://ejournal.pabki.org/index.php/ETCE>
- Baha'Uddin, Moh., & Fatah, Z. (2024). Penerapan data mining clustering K-Means dalam mengelompokkan data penduduk penyandang disabilitas. *Gudang Jurnal Multidisiplin Ilmu*, 2(10), 86–94. <https://doi.org/10.59435/gjmi.v2i11.1040>
- Fauziah, R., & Purnamasari, A. I. (2023). Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia. *Hello World Jurnal Ilmu Komputer*, 2(1), 34-41. <https://doi.org/10.56211/helloworld.v2i1.232>
- Faqih, M. F. A., & Mailoa, E. (2025). clustering zonasi daerah rawan bencana alam Provinsi Jawa Tengah menggunakan algoritma k-means dan library geopandas. *IT-EXPLORE: Jurnal Penerapan Teknologi Informasi Dan Komunikasi*, 4(1), 116–127. <https://doi.org/10.24246/itexplore.v4i1.2025.pp116-127>
- Halik, M. F. Al, & Septiana, L. (2022). Analisa data untuk prediksi daerah rawan bencana alam di Jawa Barat menggunakan algoritma K-Means clustering. *Journal of Information System, Applied, Management, Accounting and Research (JISAMAR)*, 6(4), 856–870. <https://doi.org/10.52362/jisamar.v6i4.939>
- Hasdyna, N., Dinata, R. K., & Yafis, B. (2025). Optimizing K-Means Algorithm Using the Purity Method for Clustering Oil Palm Producing Regions. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 10(1), 1-15.
- Hendrastuty, N. (2024). Penerapan data mining menggunakan algoritma K-Means clustering dalam evaluasi hasil pembelajaran siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (JIMA-ILKOM)*, 3(1), 46–56. <https://doi.org/10.58602/jima-ilkom.v3i1.26>

- Noval, M., Windarsyah, & Marleny, F. D. (2025). Implementasi algoritma K-Means untuk analisis pola penjualan pada Toko Monisa. *Jurnal Media Informatika (JUMIN)*, 6(3), 1996–2002. <http://ejournal.sisfokomtek.org/index.php/jumin>
- Nugroho, M. R., Hendrawan, I. E., & Puwantoro. (2022). Penerapan algoritma K-Means untuk klasterisasi data obat pada Rumah Sakit ASRI. *Jurnal Nuansa Informatika*, 16(1), 125–133. <https://journal.uniku.ac.id/index.php/ilkom>
- Prastiwi, H., Pricilia, J., & Raswir, E. (2022). *Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM) Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means clustering*.
- Pribadi, W. W., Yunus, A., & Wiguna, A. S. (2022). Perbandingan Metode K-Means Euclidean Distance dan Manhattan Distance Pada Penentuan Zonasi Covid-19 di Kabupaten Malang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), 493-500.
- Purba, R. F., Panjaitan, A. A., Butar-Butar, L. T., Sitorus, J. F. E. D. L., Rumapea, R., & S, I. M. S. (2024). Implementasi K-Means clustering untuk menentukan tingkat bencana rawan banjir di wilayah Sumatera Utara. *TAMIKA: Jurnal Tugas Akhir Manajemen Informatika & Komputerisasi Akuntansi*, 4(2), 210–215. [https://doi.org/10.46880/tamika.Vol4No2\(SEMNASTIK\).pp210-215](https://doi.org/10.46880/tamika.Vol4No2(SEMNASTIK).pp210-215)
- Putri, S. A. (2025). Pengelompokan daerah rawan bencana di Indonesia menggunakan metode clustering K-Means. *Jupiter: Publikasi Ilmu Keteknikan Industri, Teknik Elektro Dan Informatika*, 3(1), 9–16. <https://doi.org/10.61132/jupiter.v3i1.644>
- Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang. *Jurnal Janitra Informatika Dan Sistem Informasi*, 2(1), 9–19. <https://doi.org/10.25008/janitra.v2i1.142>
- Taryana, A., Mahmudi, M. R. El, & Becti, H. (2022). Analisis kesiapsiagaan bencana banjir di Jakarta. *JANE: Jurnal Administrasi Negara*, 13(2), 302–311.
- Virantika, E., Kusnawi, & Ipmawati, J. (2022). Evaluasi Hasil Pengujian Tingkat Clusterisasi Penerapan Metode K-Means Dalam Menentukan Tingkat Penyebaran Covid-19 di Indonesia. *Jurnal Media Informatika Budidarma*, 6(3), 1657-1666. <https://doi.org/10.30865/mib.v6i3.4325>
- Wathoriq, T. (2023). Implementasi data mining clustering berbasis web untuk mengelompokkan wilayah rawan banjir berdasarkan dampak yang terjadi dengan algoritme K-Means pada BPBD Kota Tangerang [Skripsi, Universitas Budi Luhur]. Fakultas Teknologi Informasi.

Wathoriq, T., & Subandi. (2023). Implementasi algoritme K-Means clustering untuk pengelompokkan wilayah rawan banjir. *3rd Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 2(2), 1086–1093.
<https://senafiti.budiluhur.ac.id/index.php/senafiti/index>