A GeoLLM Framework for Multimodal Data Integration in the Development of Jakarta's Smart City System

Program Riset PRIMA ITB 2026 · PPMI

Resy Meilani, S.T., M.T. [1], Prof. Dr. Albertus Deliar, S.T., M.T. [2], Dr. Budhy Soeksmantono, S.T., M.T. [2], Ir. Agung Budi Harto, M.Sc., Ph.D. [2]

[1] Program Studi Doktor Teknik Geodesi dan Geomatika, Fakultas Ilmu dan Teknologi Kebumian, ITB

[2] Kelompok Keahlian Sains dan Teknologi Informasi Geografis, Fakultas Ilmu dan Teknologi Kebumian, ITB

Abstrak

Model bahasa besar (LLM) memungkinkan analisis geospasial dijalankan dari pertanyaan berbahasa sehari-hari, tetapi kebenaran dan konsistensi jawabannya belum teruji. Penelitian ini mengukur kebenaran dan konsistensi LLM sebagai tahap awal kerangka GeoLLM untuk kota cerdas Jakarta. Claude, GPT, dan DeepSeek diuji pada analisis akses sekolah dasar di Jakarta Utara melalui 2.070 percobaan. Setiap pertanyaan diajukan lima kali, lalu jawabannya dibandingkan dengan jawaban acuan. Pada pertanyaan rinci, 81–88% jawaban benar. Pada pertanyaan ambigu, hasil antarpercobaan hanya sama pada 10–26% pasangan, dan model yang sama dapat memberi kesimpulan berlawanan. Satu model juga mengarang data ketika data tidak tersedia. Karena itu, jawaban LLM perlu diuji berulang sebelum dipakai untuk mengambil keputusan.

Kata kunci: GeoLLM, SIG otonom, kota cerdas, LLM

Pendahuluan

Mengolah data kota menjadi hasil analisis geospasial, seperti peta dan tabel, memerlukan beberapa langkah di perangkat lunak SIG. LLM dapat mengotomatiskan langkah tersebut dari pertanyaan berbahasa alami (Li & Ning, 2023; Zhang dkk., 2024).



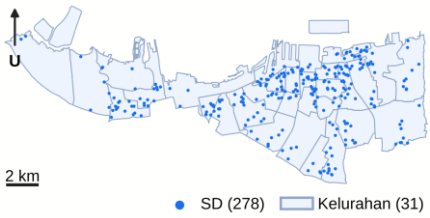
Tiga hal yang diuji:

Benar?
Apakah jawabannya sesuai jawaban acuan?

Konsisten?
Apakah lima percobaan hasilnya sama?

Jujur?
Apakah LLM menolak bila data tidak ada?

Wilayah uji: Jakarta Utara



Data: batas kelurahan, titik SD, penduduk, jaringan jalan OSM

Pustaka Utama

Li, Z., & Ning, H. (2023). Autonomous GIS: The next-generation AI-powered GIS. International Journal of Digital Earth, 16(2), 4668–4686.

Zhang, Y., Wei, C., He, Z., & Yu, W. (2024). GeoGPT: An assistant for understanding and processing geospatial tasks. International Journal of Applied Earth Observation and Geoinformation, 131, 103976.

Metode Penelitian



Rancangan uji

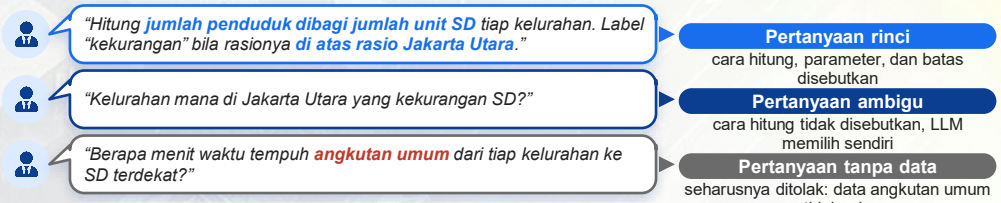
3 LLM × 23 per-tanyaan × 2 bahasa × 3 gaya kalimat × 5 kali = **2.070** percobaan

Versi: Claude Haiku 5.5, GPT-6 Luna, DeepSeek Flash

Claude · GPT · DeepSeek

Penilaian: jawaban benar bila ≥ 95% baris sama dengan jawaban acuan yang dikerjakan manual di perangkat lunak SIG dan dianggap benar.

Tiga jenis pertanyaan *pertanyaan rinci dan ambigu menanyakan hal yang sama*



Hasil dan Pembahasan

Pertanyaan rinci

81–88%

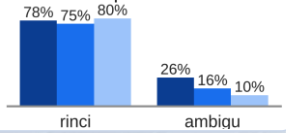
percobaan menjawab dengan benar



Pertanyaan ambigu

10–26%

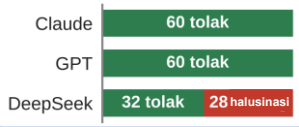
hasil sama antarpercobaan



Pertanyaan tanpa data

28 dari 60

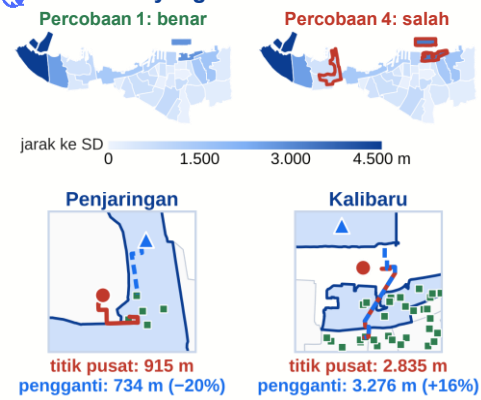
percobaan berisi data karangan



Variasi jawaban LLM terhadap pertanyaan dan model yang sama

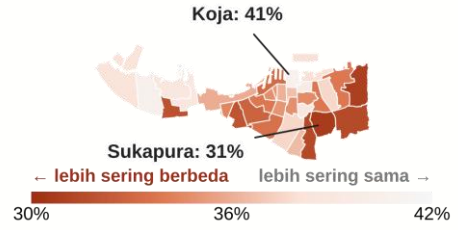


Peta salah yang terlihat benar



Titik pusat (centroid) kedua kelurahan berada di luar wilayahnya. DeepSeek diam-diam memindahkannya, sehingga jarak ke SD berubah 16–20%, padahal petanya terlihat sama.

Di mana jawaban paling sering berbeda?



Kelurahan dengan jawaban yang sering berbeda cenderung mengelompok, terutama di bagian timur (Moran's I = 0,22; p = 0,03).

Kesimpulan

1. Pada pertanyaan rinci, LLM sudah cukup tepat: 81–88% jawabannya benar.
2. Pada pertanyaan ambigu, jawaban LLM berubah-ubah antarpercobaan.
3. Sebagian kesalahan tidak terlihat di peta. Jawaban LLM perlu diuji berulang sebelum dipakai.