

Estimasi Biaya Operasional

Opsi 1: Menggunakan Gemini Flash-8B (API BASED)

| Komponen | Spesifikasi | Biaya/Bulan |
|--------------------------|---|------------------|
| Hosting Server | (Vultr Cloud Compute) 2 vCPU, 4 GB RAM, 50 GB SSD, 1 TB bandwidth | Rp. 320.000 |
| Layanan AI | (Gemini Flash-8B API) ±100 pengguna/hari, ±30.000 chat/bulan, response <2 detik | Rp. 18.000 |
| Total Biaya Operasioanal | | Rp 338.000/bulan |

Biaya Development (One-time):

- Domain (.id atau .ac.id): Rp 150.000/tahun
- SSL Certificate: Gratis (Let's Encrypt)
- Total Initial Investment: Rp 150.000

Total Biaya 4 Bulan (Periode Proyek):

- Biaya Operasional: $\text{Rp } 338.000 \times 4 = \text{Rp } 1.352.000$
- Biaya Development: Rp 150.000
- **Grand Total: Rp 1.502.000****

Opsi 2 : Menggunakan GPT-4o mini (API BASED)

| Komponen | Spesifikasi | Biaya/Bulan |
|--------------------------|---|------------------|
| Hosting Server | (Vultr Cloud Compute) 2 vCPU, 4 GB RAM, 50 GB SSD, 1 TB bandwidth | Rp. 320.000 |
| Layanan AI | (GPT-4o mini API) ±100 pengguna/hari, ±30.000 chat/bulan, response <2 detik | Rp. 25.000 |
| Total Biaya Operasioanal | | Rp 345.000/bulan |

****Biaya Development (One-time):****

- Domain (.id atau .ac.id): Rp 150.000/tahun

- SSL Certificate: Gratis (Let's Encrypt)
- Total Initial Investment: Rp 150.000

****Total Biaya 4 Bulan (Periode Proyek):****

- Biaya Operasional: $\text{Rp } 345.000 \times 4 = \text{Rp } 1.380.000$
- Biaya Development: Rp 150.000
- ****Grand Total: Rp 1.530.000****

Opsi 3 : Menggunakan GLM-4.6 (SELF-HOSTED)

| Komponen | Spesifikasi | Biaya/Bulan |
|---------------------------------|---|---------------------|
| Hosting Server | (Vultr Cloud GPU Instance) 8 vCPU, 32 GB RAM, 200 GB SSD, GPU NVIDIA A100 40GB, 2 TB bandwidth | Rp. 6.000.000 |
| Layanan AI | (GLM-4.6 Model Open Source) ± 100 pengguna/hari, ± 30.000 chat/bulan, response <2 detik (dengan optimasi GPU) | Rp. 0 |
| Total Biaya Operasioanal | | Rp. 6.000.000/bulan |

****Biaya Development (One-time):****

- Domain (.id atau .ac.id): Rp 150.000/tahun
- SSL Certificate: Gratis (Let's Encrypt)
- Total Initial Investment: Rp 150.000

****Total Biaya 4 Bulan (Periode Proyek):****

- Biaya Operasional: $\text{Rp } 6.000.000 \times 4 = \text{Rp } 24.000.000$
- Biaya Development: Rp 150.000
- ****Grand Total: Rp 24.150.000****

Kami sudah menyusun estimasi biaya operasional untuk proyek chatbot ini dengan membandingkan tiga opsi model AI yang bisa digunakan:

1. Gemini Flash-8B (API-based)

- Biaya operasional sekitar Rp 338.000/bulan.
- Lebih ekonomis, performa sudah cukup baik untuk skala kampus.
- Cocok jika prioritas utama adalah biaya yang rendah dan stabilitas layanan.

2. GPT-4o mini (API-based)

- Biaya operasional sekitar Rp 345.000/bulan.
- Jawaban model umumnya lebih natural dan akurat dibanding Gemini.
- Cocok jika klien menginginkan kualitas interaksi yang lebih baik, dengan biaya sedikit lebih tinggi.

3. GLM-4.6 (Self-hosted)

- Membutuhkan server GPU khusus, dengan biaya sekitar Rp 6.000.000/bulan.
- Sangat fleksibel, bisa dikustomisasi penuh (misalnya integrasi dokumen internal, PDF handling, RAG).
- Namun, biaya jauh lebih tinggi dan kompleksitas teknis juga lebih besar.

Kesimpulan:

- Untuk proyek akademik dengan durasi 4 bulan, opsi paling realistis adalah API-based (Gemini atau GPT-4o mini) karena lebih murah dan sederhana.
- Jika di kemudian hari dibutuhkan fleksibilitas lebih tinggi (misalnya integrasi data kampus skala besar), maka opsi self-hosted (GLM-4.6) bisa dipertimbangkan.