Part 1 — Cost Research

Objective

Understand how Google’s Gemini (or any LLM) usage is priced.

Estimate costs for API calls during a hackathon project.

Steps

Identify Model Versions & Rates

Check different versions of the LLM and note their cost per 1,000 tokens or per API request.

Example:

Model Billing Unit Cost per Unit Notes

Gemini Alpha per 1k tokens $0.03 Best for medium-length prompts

Gemini Beta per 1k tokens $0.08 Handles longer prompts, more detailed

Gemini Free per 1k tokens Free (up to 15k tokens/month) Ideal for testing/prototyping

Check Free / Limited-use Plans

Many models offer limited free tokens per month to test.

Useful for initial experiments without spending money.

Understand Token Usage & Cost

Tokens = pieces of text (roughly 1 token ≈ 4 characters).

Longer prompts or longer AI responses increase cost.

Calculate Sample Costs

Example 1: 500-token prompt + 500-token response using Gemini Alpha

Total tokens = 1000

Cost = 1000 / 1000 \* $0.03 = $0.03 per API call

Example 2: 2,000-token prompt + 1,000-token response using Gemini Beta

Total tokens = 3,000

Cost = 3,000 / 1000 \* $0.08 = $0.24 per API call

Outcome:

Teams can estimate the budget for LLM usage, decide which model version is cost-effective, and plan prompt length accordingly