Assignment 1 - Evaluating the Cost of AI Systems

I imagined implementing AI as a customer support assistant for Singtel in Singapore, where I work. The AI would handle frequently asked questions, such as roaming activation, billing inquiries, plan upgrades, and data usage monitoring.

For this exercise, I used the prompt:
 "I am travelling from Singapore and I am using Singtel postpaid. How do I activate international roaming on my mobile plan? Will I be charged upfront, or only when I start using data? What are the available roaming packages for Asia and Europe? Is there a way to monitor my roaming usage in real-time?"

I measured the token usage using OpenAI's Tokenizer. The input was approximately **45 tokens**, and the model's response was around **360 tokens**. I experimented with different models **Model type:** GPT-4 vs GPT-4o-mini and **Number of requests:** 50,000

The results showed that costs vary dramatically depending on the model and response length. For 50,000 queries, GPT-4 would cost approximately **$4,575**, while GPT-4o-mini would cost only **$45.38**. This highlights the importance of balancing quality with cost, especially for high-volume repeated queries.

Although I am **not using caching in this example**, LLM token caching is an important optimization technique. It stores and reuses parts of processed prompts (such as system instructions or documents) or intermediate calculations (KV cache) to avoid repetitive computation.

Below is my observation

Prompt: I am travelling from Singapore and I am using Singtel postpaid. How do I activate international roaming on my mobile plan? Will I be charged upfront, or only when I start using data? What are the available roaming packages for Europe?

Input Tokens:48

I am travelling from Singapore and I am using Singtel postpaid. How do I activate international roaming on my mobile plan? Will I be charged upfront, or only when I start using data? What are the available roaming packages for Europe?

Clear | Show example

Tokens
48

Characters
234

I am travelling from Singapore and I am using Singtel postpaid. How do I activate international roaming on my mobile plan? Will I be charged upfront, or only when I start using data? What are the available roaming packages for Europe?

Text | Token IDs

Side-by-side look at gpt-4o-mini and gpt-4 using one prompt

gpt-4o-mini produced 268 output tokens, Generated below response:

To activate international roaming on your Singtel postpaid plan while traveling from Singapore, follow these steps:

Activate International Roaming:

You can activate international roaming through the My Singtel app, the Singtel website, or by calling customer service.

If you're using the app, log in to your account, navigate to "Roaming," and follow the prompts to activate it.

Charges:

Typically, you are not charged any upfront fee for activating international roaming. Charges will only apply when you start using data, making calls, or sending texts while overseas. It's advisable to check for any specific terms related to your plan.

Roaming Packages for Europe:

Singtel usually offers various roaming packages for Europe. Common options include daily or weekly plans that provide data, local calls, and SMS.

As packages may change, it is best to check the Singtel website or the My Singtel app for the latest offerings and pricing. Options may include plans like "Data Roaming Passes" or "Roam Like Home."

Pre-Purchase and Check:

Consider purchasing a roaming package before you travel to manage costs effectively. You can also check your current usage and remaining balance on the app while abroad.

If you have specific preferences for data usage or need assistance with activating roaming, it might be best to contact Singtel's customer service directly for personalized help.

To activate international roaming on your Singtel postpaid plan while traveling from Singapore, follow these steps:
Activate International Roaming:
You can activate international roaming through the My Singtel app, the Singtel website, or by calling customer service.
If you're using the app, log in to your account, navigate to "Roaming," and follow the prompts to activate it.
Charges:
Typically, you are not charged any upfront fee for activating international roaming. Charges will only apply when you start using data, making calls, or sending texts while overseas. It's advisable to check for

Clear   Show example

Tokens   **Characters**
268   1385

To activate international roaming on your Singtel postpaid plan while traveling from Singapore, follow these steps:
Activate International Roaming:
You can activate international roaming through the My Singtel app, the Singtel website, or by calling customer service.
If you're using the app, log in to your account, navigate to "Roaming," and follow the prompts to activate it.
Charges:
Typically, you are not charged any upfront fee for activating international roaming. Charges will only apply when you start using data, making calls, or sending texts while overseas. It's advisable to check

---

gpt-4 produced 336 output tokens, Generated below response:

You can activate international roaming on your Singtel postpaid plan by following these steps:

Log in to My Singtel app.

Under the Home setting, select 'Roaming'.

Toggle 'Pay-As-U-Roam' to activate.

There's no subscription charge for Pay-As-U-Roam. It will only be a charge-as-you-use service.

For roaming in Europe, Singtel offers the 'DataRoam Saver Plan' (EU) which you can subscribe for daily unlimited data usage in 43 European countries. The price per day can vary, you can check at the Singtel website or app for the current rate.

Also, Singtel offers Roaming Plans. These come in:

a. 'ReadyRoam 1GB 30 Days' – Activate this for 1GB data, valid 30 days, in 56 Destinations across the Americas, Asia, Europe, Middle East and Oceania.

b. 'ReadyRoam 500 MB 7 Days' – Activate this for 500MB data, valid 7 days, in 16 Destinations across Asia and Oceania only.

Please note rates are subject to changes and may vary depending on the country of usage. It is advisable to double check the rates and terms on the Singtel website or app before you travel.

Remember to enable data roaming on your mobile device under your phone settings only when you are ready to use the service. Data charges will start once you begin using your data abroad, based on the plan you subscribed to. If you do not wish to use data at any point, turn off data roaming on your mobile device to prevent additional charges.

You can use the tool below to understand how a piece of text might be tokenized by a language model, and the total count of tokens in that piece of text.

GPT-4o & GPT-4o mini | **GPT-3.5 & GPT-4** | GPT-3 (Legacy)

You can activate international roaming on your Singtel postpaid plan by following these steps:
Log in to My Singtel app.
Under the Home setting, select 'Roaming'.
Toggle 'Pay-As-U-Roam' to activate.
There's no subscription charge for Pay-As-U-Roam. It will only be a charge-as-you-use service.
For roaming in Europe, Singtel offers the 'DataRoam Saver Plan' (EU) which you can subscribe for daily unlimited data usage in 43 European countries. The price per day can vary, you can check at the Singtel website or app for the current rate.
Also, Singtel offers Roaming Plans. These come in:
a. 'ReadyRoam 1GB 30 Days' – Activate this for 1GB data, valid 30 days, in 56 Destinations across the

Clear | Show example

Tokens
336

Characters
1397

You can activate international roaming on your Singtel postpaid plan by
 following these steps:
Log in to My Singtel app.
Under the Home setting, select 'Roaming'.
Toggle 'Pay-As-U-Roam' to activate.
There's no subscription charge for Pay-As-U-Roam. It will only be a

Cost:

helicone | Docs | Pricing | Models | Resources ∨ | Tools ∨ | Careers | 4.8K | Contact us | Log In

LLM API Pricing Calculator

Calculate the cost of using AI models with Helicone's free pricing tool.

Input Tokens 50   Output Tokens 1500   Share

Search models   Filter by Provider   Filter by Model   Clear All

OpenAI ⊗

gpt-4 ⊗   gpt-4o-mini ⊗

PROVIDER ∨	MODEL ∨	INPUT/1K TOKENS ∨	OUTPUT/1K TOKENS ∨	INPUT COST ∨	OUTPUT COST ∨	TOTAL COST ∨
OpenAI	gpt-4	$0.03	$0.06	$0.0015	$0.09	$0.0915
OpenAI	gpt-4o-mini	$0.00015	$0.0006	$0.0000075	$0.0009	$0.0009075

Provider	Model	Input Cost / Request	Output Cost / Request	Total Cost / Request	Number of Requests	Total Cost
OpenAI	GPT-4	$0.0015	$0.09	$0.0915	50,000	$4,575
OpenAI	GPT-4o-mini	$0.0000075	$0.0009	$0.0009075	50,000	$45.38

Note: Token counts differ across LLMs because each model uses its own **tokenizer**, which splits text into pieces according to unique rules and vocabularies. For example, OpenAI's GPT models use a variant of **byte-pair encoding (BPE)**, Claude uses a **SentencePiece BPE tokenizer**, and Google Gemini uses a **SentencePiece Unigram tokenizer**. These tokenizers segment text differently depending on their training corpus, merge rules, and vocabulary, so the same sentence can produce more tokens in one model and fewer in another.