

Hackaton ByteMe

Idei

1. frontend: platforma sa te poti conecta cu numarul de telefon si acolo sa ai un dashboard cu toate companiile la care ai apelat + diferite documente salvate special pentru tine
2. backend
 - asr automatic speech recogniziton - whisper e gratis
 - rag (chroma db pentru a stoca embedding-uri la diferite documente)
 - tts (open ai tts , cel mai ieftin si sigur)
3. **Tooluri** posibile utile
 - Live kit ([LINK](#) real time communication)
 - Plivo ([LINK](#) sms api + voice api)

Mentori

1. Putem doar sa integram tot apelul cu toate intrebarile si dupa sa primeasca pe telefon un sms ca un link de la site-ul nostru
2. Document cu detaliile avute din conversatie, pe care un broker ar putea sa ,generare de raport pe baza de conversatie, o lista sa fie sunati pentru detalii
3. Facem un flow foarte specific, sa zica la final o campanie, sa gandim mereu ce imbunatatiri putem avea
4. Voice log-uri pe backend, daca s-o blocat unde si de ce?

Subscription Plans

1. **0–10k tokens** — Free -> cheapest model : gpt-3.5-turbo-0125 *our costs are minimal because 2m tokens (1m input + 1m output) are 2 dollars. Comparison:* a small call center (e.g. a construction firm with only 1 person handling calls) spends around **1k dollars/month** in salaries.
2. **10k–100k tokens** — Business -> mid range model : gpt-4o-mini *Comparison:* a medium call center with 3–5 **employees** answering calls spends around **3k–5k dollars/month** in salaries.
3. **100k+ tokens** — Enterprise / Custom *Comparison:* a large call center with **10+ employees** answering calls spends **10k+ dollars/month** in salaries.

Pay as You Go

Cheapest model (gpt-3.5-turbo-0125)

- **1m tokens (input):** our cost is 0.5 dollars → we charge **50 dollars**
- **1m tokens (output):** our cost is 1.5 dollars → we charge **100 dollars**

Comparison : a small calling center (e.g. a construction firm has only 1 person that responds to calls :)

Mid-range model (gpt-4o-mini)

- **1m tokens (input):** our cost is 3 dollars → we charge **100 dollars**
- **1m tokens (output):** our cost is 6 dollars → we charge **200 dollars**

Expensive model (gpt-4o)

- **1m tokens (input):** our cost is 5 dollars → we charge **200 dollars**
- **1m tokens (output):** our cost is 15 dollars → we charge **400 dollars**

Why use our service?

By using our service you can reduce operational costs by 70–90%, while offering 24/7 availability and scalable support.

- Example 1: A small firm that spends **1,000 USD/month** on one call agent could replace most interactions with AI at a cost of only **50–100 USD/month** in tokens.
- Example 2: A medium call center with **5 agents** costing around **5,000 USD/month** could reduce costs to under **300 USD/month** in AI usage.
- Example 3: A large call center with **10+ agents** spending **10,000 USD or more** monthly could reduce costs to approximately **600–1,000 USD/month** with AI.
- Scalability: Unlike human agents, the AI can handle **hundreds of calls simultaneously**, so there's no need to hire extra staff during peak hours.
- ROI: Even a conservative adoption can save thousands of dollars per month, while improving response time and consistency.