Caprae Capital Partner Assessment

V. Viswha

GitHub Repo Link: https://github.com/viswha2012/Caprae_Capital_Partner_Leads_Tool_.git
YouTube Demo Link: https://youtu.be/r3W6yBv2phA

The leads application is designed and developed to streamline lead generation process in the aim of making it a one stop platform for all Private Equity and MaaS Services. The system follows a 3-step workflow:

1. **Lead Discovery**: Fetch basic company information from Google Maps API
2. **Lead Enrichment**: Augment data with Apollo.io's organization enrichment API
3. **Outreach Generation**: Use Mistral AI to create personalized outreach emails

Approach:

- **User Input Layer:** Users enter free-form queries specifying the type of companies they are looking for (e.g., "Health Tech startups in London with Series A funding")
- **LLM Based Query Parser:** We use a local LLM (mistral 7B model running via Ollama) to parse this input and output structured query parameters compatible with our Apollo API.
- **Data Enrichment:** Using the Apollo API and crawl4ai, we enrich each company with metadata such as funding rounds, technology stack, LinkedIn URLs, and company descriptions.
- **Outreach Automation:** For each enriched company, an AI-generated personalized email is created suggesting a meeting to explore M&A or investment opportunities.

Model Selection: Mistral 7B via Ollama
Reasoning: Mistral is lightweight and efficient for local inference. It displays strong performance in task specific prompt-following. Its open source and idea for secure enterprise deployments. Provides accurate information extraction and professional text generation.

Data Pre Processing:

- **Query Parsing:** The prompt is templated with a consistent schema so the model can reliably output a Python dictionary with search parameters.
- **Website Crawling:** Using crawl4ai, we extract readable text from company websites to inform the email generation prompt.

Performance:

- **Accuracy (Manual Checks):** ~90% of parameter extraction from user queries was relevant and correctly formatted.
- **Email Quality:** Emails were rated "professional" and "clear" in tone in user evaluations.
- **Latency**: Each enrichment + email generation takes ~4–7 seconds per company on a local enabled CPU.