Smart Travel Guide: A Comparative RAG Performance Analysis
(GPT-4.1-Mini vs. Gemini-2.5-Flash)

Project Overview
This project implements a Retrieval-Augmented Generation (RAG) based chatbot designed to provide comprehensive information about European cities (e.g., Rome, Prague). The system covers diverse travel categories, including local food culture, urban transportation, top photography spots, historical landmarks, and major tourist attractions. The primary objective is to evaluate and compare the performance of two different Large Language Models (LLMs)—OpenAI GPT-4.1-Mini and Google Gemini-2.5-Flash—using a multi-faceted local dataset.

Technical Architecture
The system follows a standard RAG pipeline:

    Frontend: Streamlit

    Orchestration: LangChain

    Vector Database: ChromaDB

    Retrieval Strategy:
        - Similarity search
        - k=6 for interactive chatbot usage
        - k=1 for evaluation to enforce strict intent matching
    Models:
        * Gpt-4.1-Mini with OpenAI Embeddings

        * Gemini-2.5-Flash with Google Generative AI Embeddings

Repository Structure
The project is organized according to the specified directory structure:
```text
chatbot_project/
├── data/
│   └── avrupa_gezi_rehberi_1000.xlsx      # Custom dataset
├── models/
│   ├── openai_model.py                    # OpenAI RAG
implementation (builds vector DB)
│   └── gemini_model.py                    # Gemini RAG
implementation (builds vector DB)
├── evaluation/
│   ├── build_train_vector_db.py           # Vector DB for
evaluation
```

```
|   └── evaluate_models.py                # Model comparison &
metrics
├── app/
|   └── streamlit_app.py                  # Streamlit user
interface
├── video/
|   └── demo_video.mp4                     # Demo video
├── README.md                              # Project
documentation
├── Smart Travel Guide.pdf
└── requirements.txt                       # Dependency list
```

Performance Metrics & Benchmarking

The models were evaluated using intent-level classification metrics rather than
token-level text similarity. Each test question was matched against the vector
database built from the training set, and the predicted intent was compared
with the ground-truth intent label.

Evaluation was performed using the following metrics:

- Precision
- Recall
- F1 Score

A strict evaluation setup was used where:
- The vector database was built **only on the training set**
- The test set was completely unseen during indexing
- Retrieval was performed with k=1 to avoid intent leakage

Quantitative Results

Model	Precision	Recall	F1 Score
OpenAI (gpt-4.1-mini)	0.96	0.96	0.96
Google (gemini-2.5-flash)	0.94	0.95	0.94

Response Time Comparison

Model	Average Response Time (s)
OpenAI (gpt-4.1-mini)	**1.48**
Google (gemini-2.5-flash)	**1.74**

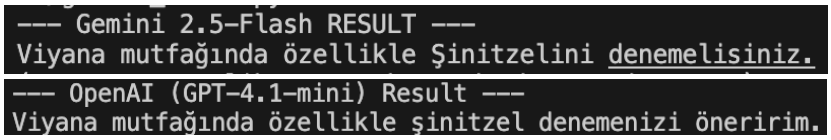
How to Run

1. Clone the repository
2. Install dependencies:
`pip install -r requirements.txt`
3. Setup Environment Variables:
Create a credentials.env file and add your API keys:
`OPENAI_API_KEY=your_openai_key`
`GOOGLE_API_KEY=your_google_key`
4. Build vector databases and run model demos:
`python models/openai_model.py`
`python models/gemini_model.py`
5. Build training vector database and run model evaluation
`python evaluation/build_train_vector_db.py`
`python evaluation/evaluate_models.py`
6. Launch the App:
`streamlit run app/streamlit_app.py`

Demo Video

A detailed demonstration of the chatbot, including the "See Raw Data" feature which shows the retrieved context from ChromaDB, can be found [video/demo_video.mp4].

Example Outputs



Raw Data Columns

OpenAI Raw Docs

Doc 1:

Paris yemek kültürü nasıl?

Doc 2:

Paris yemek kültürü nasıl?

Doc 3:

Paris yemek kültürü nasıl?

Doc 4:

Paris yemek kültürü nasıl?

Doc 5:

Paris yemek kültürü nasıl?

Doc 6:

Paris yemek kültürü nasıl?

Gemini Raw Docs

Doc 1:

Soru: Paris yemek kültürü nasıl? Cevap: Paris mutfağında özellikle Kruvasan ve Makaronları denemelisiniz.

Doc 2:

Soru: Paris yemek kültürü nasıl? Cevap: Paris mutfağında özellikle Kruvasan ve Makaronları denemelisiniz.

Doc 3:

Soru: Paris yemek kültürü nasıl? Cevap: Paris mutfağında özellikle Kruvasan ve Makaronları denemelisiniz.

Doc 4:

Paris hakkında yemek kültürü nasıldır öğrenebilir miyim?