SC Onboarding Agent: Reducing Customer Support Load with AI-Powered Assistance

Introduction

This project was created for the 'Practical Application of Artificial Intelligence' course at the Faculty of Electrical Engineering, University of Sarajevo. It addresses the real-world challenge of reducing support workload in a company that manages dozens of applications. The team aimed to automate onboarding and reduce ticket volume through an AI Onboarding assistant.

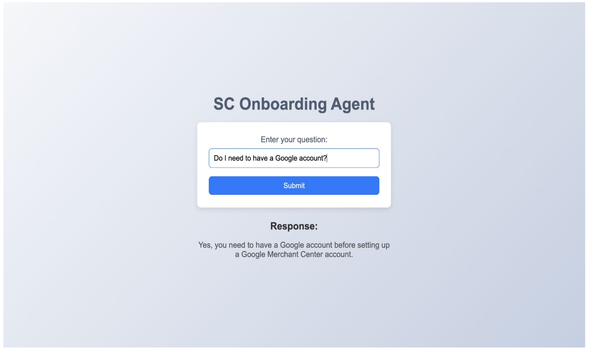
Project Goals

- Minimize incoming support tickets.
- Improve response accuracy and speed.
- Enhance the knowledge base and reduce repetitive queries.
- Support automatic/manual customer interaction modes.

Project Description

The onboarding agent integrates LangChain and LangGraph to deliver real-time document-based Q&A. Outdated manuals are removed and new ones vectorized using MistralAI embeddings, then stored in-memory. It was built for Shopify Partner apps but is adaptable to any context. The system is lightweight and includes a Python Flask-based frontend for interaction.

Frontend



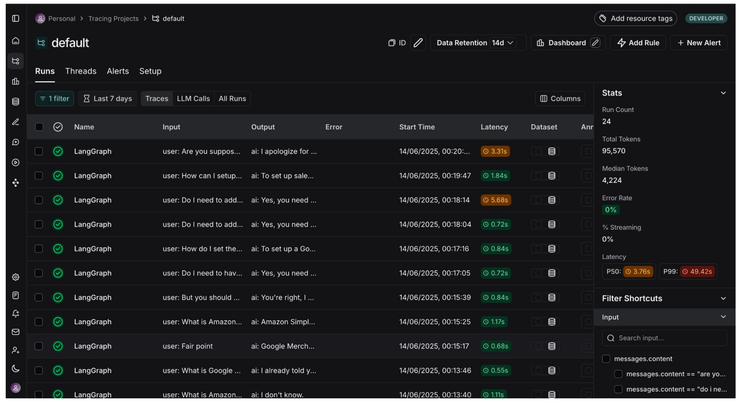
Evaluation Strategy

LangSmith's LLM-as-a-judge framework is used for evaluation. A dataset of realistic Q&A examples is assessed using LLaMA3-70B. The model validates factual correctness and identifies errors, response latency, and formatting consistency. Results are visualized in LangSmith dashboards for transparency.

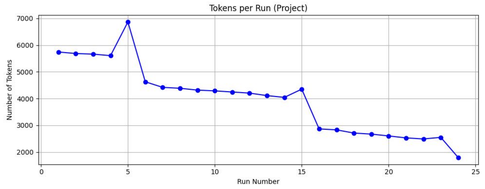
Results

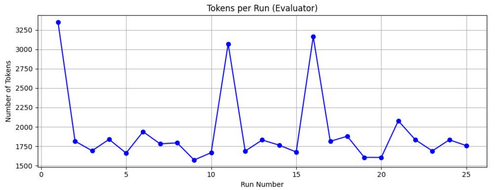
The actual project runs processed significantly more tokens per run on average, which explains the longer response times compared to the evaluators. Despite the larger workload, the project runs maintained a excellent error rate over 24 runs, demonstrating stable performance. In contrast, the evaluator runs involved smaller token counts, leading to faster responses but a higher error rate of 8%, likely due to token limits and edge-case testing scenarios. Overall, the real project usage shows consistent reliability under heavier loads, while the evaluators highlight areas where errors may occur under constrained conditions. See readme.md file for details.

LangSmith Evaluation Dashboard



Token usage





References

- LangChain & LangGraph official docs
- Research on vectorized document retrieval
- LangSmith's framework for LLM evaluation
- Internal support documentation and user manuals

Conclusions & Implications

- Significant reduction in support workload
- Higher consistency in onboarding experience
- Scalable solution across applications and languages
- Sets groundwork for multilingual, memory-enhanced AI agents

Acknowledgements

- LangChain, LangGraph, LangSmith
- Groq, MistralAI, Hugging Face
- Flask, pandas, BeautifulSoup4, python-docx