E-COMMERCE CHAT ASSISTANT

CONVERSATIONAL AI FOR SMARTER SHOPPING ASSISTANCE.

Presented by Vaibhav Garg

PROBLEM STATEMENT

- Traditional e-commerce platforms often overwhelm users with excessive information, making it difficult to quickly find relevant products or get clear answers to common questions.

- This results in high bounce rates, unanswered FAQs, and dependence on slow or inefficient customer support that diminishes user satisfaction.

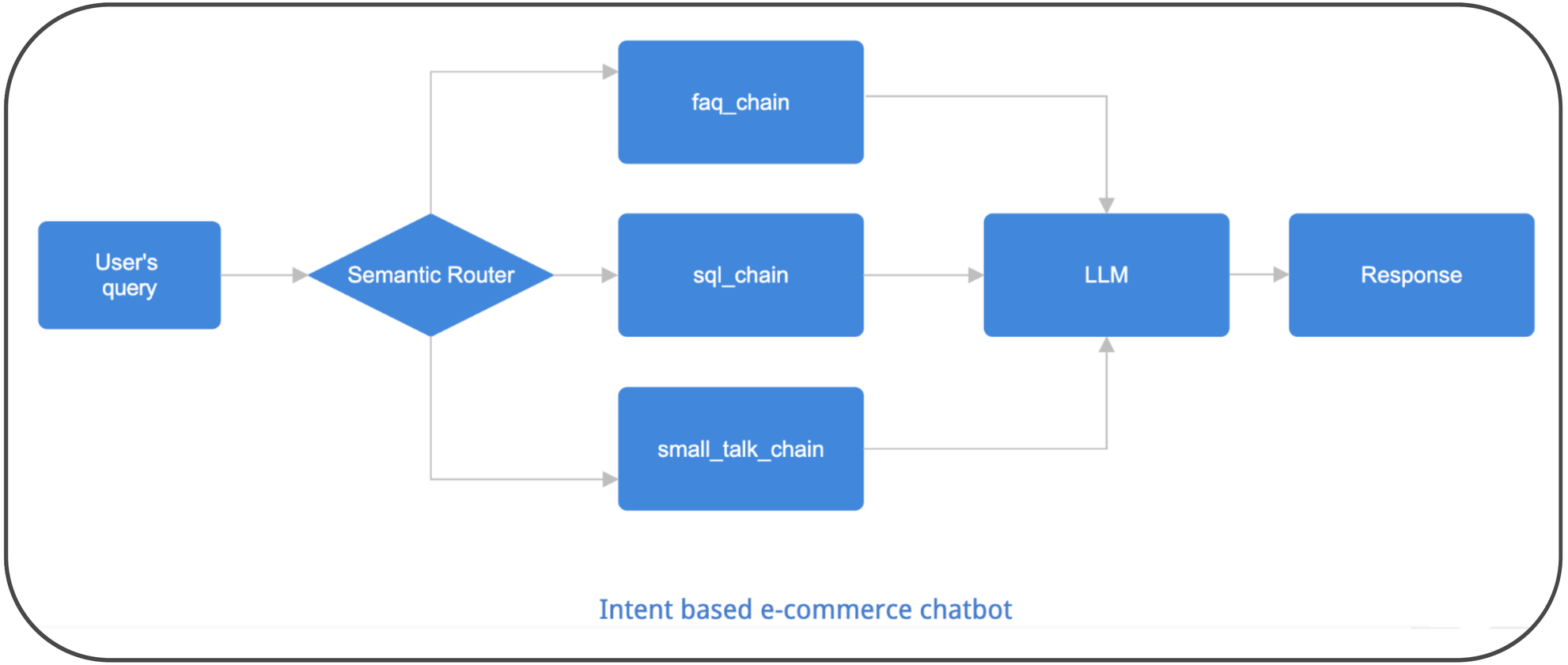
- There is a growing need for a personalized, intelligent assistant that can deliver dynamic, helpful responses — whether it's locating products, resolving policy-related queries, or engaging in small talk.

- An AI-powered conversational agent equipped with product search, FAQ resolution, and natural dialogue capabilities can streamline user experience, reduce friction, and enhance engagement on e-commerce platforms.

PROJECT OBJECTIVES

- Develop a conversational AI assistant tailored for e-commerce that intelligently understands and responds to diverse user intents.
- Handle product search queries through semantic-to-SQL translation over a structured product database.
- Retrieve and answer policy-related FAQs using vector similarity from embedded documents stored in a ChromaDB.
- Respond to small talk or general queries with friendly, natural language interactions.
- Deliver a unified, interactive experience via a Streamlit-based web interface powered by large language models and intent-aware routing logic.

SYSTEM ARCHITECTURE



TECHNOLOGY STACK

This project integrates cutting-edge AI and lightweight web technologies to deliver a seamless and intelligent conversational experience on e-commerce platforms:

- **Frontend**
  - **Streamlit :** Used to build a responsive and interactive chat interface for handling product queries, FAQs, and casual conversation.

- **Language Model**
  - **Groq + LLaMA 3.3 :** Provides ultra-fast inference for real-time natural language understanding and response generation across all user intents.

- **Semantic Routing & Embeddings**
  - **Sentence Transformers :** Uses the all-MiniLM-L6-v2 model for semantic embedding of user queries and intent classification.
  - **Semantic Router :** Determines whether the input query is a product search, FAQ, or small talk using embedding similarity

- **FAQ Retrieval**
  - **ChromaDB :** Stores embedded FAQs locally and enables semantic similarity search for answering frequently asked questions accurately.

- **Product Query Handling**
  - **SQLite :** Manages the e-commerce product catalog and enables dynamic SQL query execution for relevant product retrieval.

FUNCTIONAL WORKFLOW:
MULTI-INTENT CONVERSATIONAL AI AGENT

This chatbot system intelligently routes user queries through semantic understanding to provide instant, relevant responses across FAQs, product searches, and casual conversation — all within a unified interface.

1. Input & Preprocessing

- Users interact via a clean Streamlit UI.
- Each input is preprocessed (e.g., normalization, tokenization).
- Sentence Transformers encode the input into a semantic vector.

2. Intent Classification (Semantic Routing)

- The embedded input is compared to pre-defined intent embeddings using cosine similarity.
- The most probable intent is selected:
  - FAQ
  - Product Search (SQL)
  - Small Talk

3. Intent-Specific Processing

a. FAQ (Knowledge Base Retrieval)

- Query is used to perform a similarity search in ChromaDB (pre-embedded FAQ corpus).
- Relevant chunks are retrieved.
- A grounded response is generated using Groq LLM (LLaMA 3.3) based on the retrieved context.

b. Product Search (Dynamic SQL Execution)

- Natural language query is translated into a SQL query by the LLM.
- The SQL query is executed on the SQLite product database.
- Retrieved product info is formatted into a human-readable reply by the LLM.

c. Small Talk (Direct LLM Response)

- Casual conversation is directly handled by the LLM.
- No database or retrieval step involved.

4. Response Generation & Delivery

- Responses are grounded in actual product data, FAQs, or retrieval context.
- Groq-hosted LLaMA 3.3 ensures coherent, fluent replies.
- Multi-intent support ensures smooth, consistent interactions.
- Streamlit UI delivers responses in real time with a human-like tone.

KEY FEATURES

- **End-to-End Conversational Agent**
  - From understanding user intent to delivering precise answers, the system automates the full conversational flow across FAQs, product queries, and small talk.

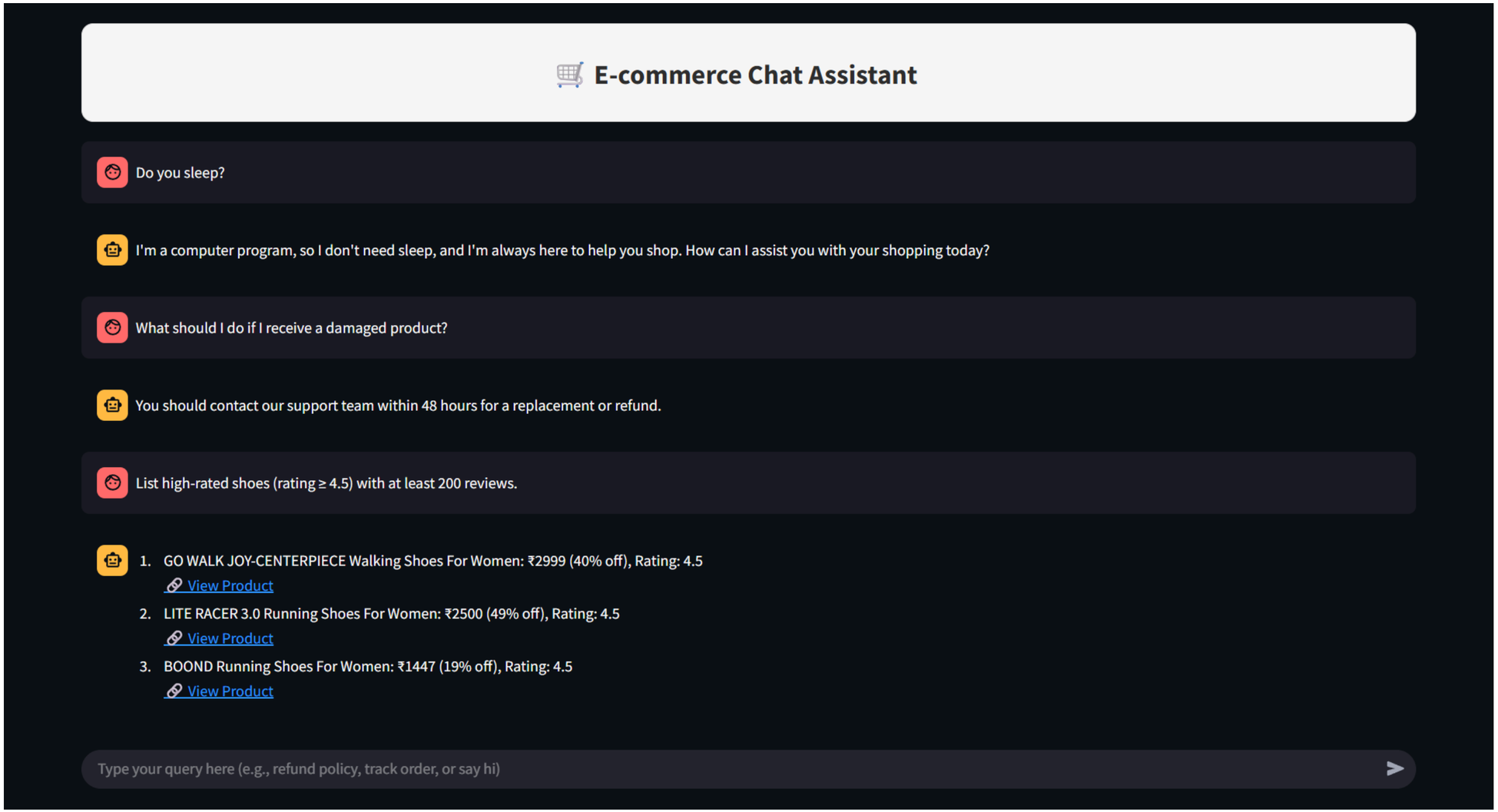
- **Natural Language Q&A**
  - Users can interact using plain English, and receive accurate, context-aware responses using the LLaMA 3.3 model via Groq's ultra-fast inference engine.

- **AI-Powered Intent Routing & Retrieval**
  - Leverages Sentence Transformers for semantic intent classification and ChromaDB for retrieving the most relevant FAQ content, enabling personalized and intelligent responses.

- **Streamlit-Powered Web UI**
  - An interactive, easy-to-use chat interface allows users to engage naturally with the assistant, offering a seamless and responsive experience.

- **Modular & Scalable Architecture**
  - Designed with separation of concerns—intent classification, response generation, and data access—making it easily extendable to new domains or functionality.

USER INTERACTION PREVIEW



PROJECT SUMMARY

- Built a conversational AI agent for e-commerce that understands diverse user intents — including FAQs, product search, and casual small talk — via a unified interface.

- Preprocessed FAQs and product data; stored FAQ embeddings in ChromaDB and product listings in a dynamic SQLite database.

- Used Sentence Transformers for semantic intent classification and Groq's LLaMA 3.3 model for ultra-fast, context-aware response generation.

- Implemented intent routing logic:
  - FAQ queries use context-grounded document retrieval from ChromaDB.
  - Product queries are converted into SQL via LLM and executed on SQLite.
  - Small talk is directly handled via LLM prompts.

- Integrated the full workflow into an interactive Streamlit app that allows users to chat naturally and receive fast, relevant answers in real time.

- Live App: https://vaibhav-project-e-commerce-chat-assistant.streamlit.app/

- GitHub Repository: https://github.com/vaibhavgarg2004/E-commerce-Chat-Assistant

THANK YOU