E-Invoice Chatbot - User Guide

Overview

The E-Invoice Chatbot is an interactive Streamlit application that allows users to ask questions about e-invoice data in both English and Arabic. The chatbot provides intelligent responses with interactive visualizations, focusing on tax compliance, fraud detection, revenue analysis, and geographic distribution.

Features

- **Bilingual Support**: Full support for both English and Arabic interfaces and responses
- **Table-Based Filtering**: Intelligently routes queries to the appropriate data tables
- **Domain-Specific Knowledge**: Focuses responses on relevant domains (tax compliance, fraud detection, etc.)
- **Interactive Visualizations**: Automatically generates appropriate charts based on query context
- **ChatGPT Integration**: Uses OpenAI's API for natural language understanding and generation

Setup Instructions

1. **Install Dependencies**: `bash pip install streamlit pandas numpy plotly openai`
2. **Run the Application**: `bash cd e_invoice_poc streamlit run chatbot/app.py`
3. **API Key Configuration**:
4. Enter your OpenAI API key in the sidebar
5. The chatbot will use this key for all API calls
6. Without a valid API key, the chatbot will use mock responses

Using the Chatbot

Language Selection

- Choose between English and Arabic using the buttons in the sidebar

- The entire UI will update to reflect your language choice

Data Filtering

- Use the sidebar dropdowns to filter by specific tables or domains
- This helps focus the chatbot's responses on relevant data

Asking Questions

- Type your question in the chat input box and press Enter or click Send
- Example questions:
- "Show me the distribution of invoices by emirate"
- "What are the most common anomaly types in invoices?"
- "Compare tax compliance rates across different sectors"
- "أظهر لي توزيع الفواتير حسب الإمارة"
- "ما هي أنواع الشذوذ الأكثر شيوعًا في الفواتير؟"

Viewing Visualizations

- Interactive charts will appear automatically for relevant queries
- Hover over chart elements to see detailed information
- Use the chart controls to zoom, pan, or download the visualization

Architecture

The chatbot is built with a modular architecture:

1. **app.py**: Main Streamlit application and UI
2. **data_router.py**: Routes queries to appropriate data tables
3. **response_handler.py**: Handles multilingual responses and domain constraints
4. **response_generator.py**: Integrates with ChatGPT API for response generation
5. **visualization_generator.py**: Creates interactive visualizations based on query context

Customization

Adding New Data Tables

To add new data tables, update the `load_data()` function in `app.py` and add corresponding keywords in `data_router.py`.

Extending Domain Knowledge

To add new domains, update the `domain_keywords` and `domain_constraints` dictionaries in the respective modules.

Adding Visualization Types

To add new visualization types, implement additional chart generation functions in `visualization_generator.py`.

Troubleshooting

- **API Key Issues**: Ensure your OpenAI API key is valid and has sufficient quota
- **Data Loading Errors**: Check that the data files exist in the expected location
- **Visualization Errors**: Verify that the data contains the expected columns for visualization

GitHub Hosting

To host this project on GitHub:

1. Create a new repository on GitHub
2. Initialize a local Git repository: `bash cd e_invoice_poc git init git add chatbot/ chatbot_design.md chatbot_readme.md git commit -m "Initial commit of e-invoice chatbot"`
3. Connect to your GitHub repository: `bash git remote add origin https://github.com/yourusername/e-invoice-chatbot.git git push -u origin main`

Deployment with Streamlit Cloud

1. Push your code to GitHub as described above
2. Go to Streamlit Cloud
3. Sign in with your GitHub account
4. Click "New app" and select your repository
5. Set the main file path to `chatbot/app.py`
6. Deploy the application