Milestone Project:

Build and deploy an Agentic AI Application to solve internal business challenges

Objective:

This milestone project provides students with the opportunity to apply their knowledge of agentic AI and the skills they have developed throughout the course to address a real-world business challenge in an enterprise setting. Students are expected to build and deploy an agentic AI application that incorporates multimodal models, agentic frameworks and architectures, Retrieval-Augmented Generation (RAG) capabilities, integration with third-party applications such as databases, and Python libraries to streamline user interface development. The application must be deployed on a server, safeguarded with appropriate guardrails, and monitored through a dashboard.

Deliverables

1. Use Case Identification Report
    a. Review some processes or procedures in your company and identify potential improvements that could be achieved using an Agentic AI application.
    b. Assess the use cases you encounter using the metrics you learned in Part 2 of this course.
    c. Design your AI agents and the tools they will require. Identify any potential retrieval systems that may be needed.
2. Document Retrieval System
    a. Users can upload documents into the application, which then runs a preprocessing step.
    b. Handle any images or tables contained within the documents.
    c. Set up and configure the vector store.
3. Agentic Workflows with LangGraph
    a. Implement agentic workflows using LangGraph to orchestrate complex tasks like document parsing, summarization, and question-answering.

    b. Enable multi-agent collaboration for tasks like fact-checking, sentiment analysis, and content generation.
4. Multimodal AI Support
    a. Integrate multimodal AI capabilities to process and analyze text, images, and combined content.
    b. Use models like CLIP (for image-text understanding) or GPT-4 Vision for multimodal tasks.
5. Production Deployment
    a. Deploy the system to a production environment (e.g., AWS, GCP, or Azure).
    b. Ensure scalability, reliability, and security.
6. User Interface
    a. Build an intuitive Streamlit or React-based UI for document upload, querying, and analysis.
    b. Include features like real-time feedback, visualizations, and interactive dashboards.
7. Guardrails and Monitoring
    a. Implement AI guardrails applied to input side and output.
    b. Set up a monitoring system with a dashboard to visualize how often agents succeed/fail, response times, number of tokens, etc.
    c. Define key metrics to be continuously monitored.

More Challenges

1. [Feedback Loop] Implement a feedback loop by creating a system that allows users to indicate whether they find the agent's responses helpful or to provide the correct answer. Use this feedback to continuously improve the AI agent's performance.
2. [Voice Command] Develop a feature that enables users to ask questions using voice input, with the application responding in both text and speech.
3. [Agent Personalization] Allow users to configure agent "personas" or select modes like "summary," "technical," or "creative."
4. [Chain-of-Thought Visualization] Display agent decision paths or reasoning chains visually

Submission Process

1. Start a new Github repo for your milestone project and start building.

2. Prepare a PowerPoint for your final presentation.
3. Practice your demo. Record a demo video in case something went wrong during the presentation.
4. **Submit your Github repo and presentation slides to the class coordinator.**
5. Submit your assignment here