Math Agent: Final Proposal

1. Input & Output Guardrails for Privacy

Approach:

- The Math Agent uses a Guardrails class to check both user input and system output for:
- Personal Identifiable Information (PII) such as emails, phone numbers, and addresses.
- Ensures only mathematics or education-related questions are processed.
- Guardrails are enforced in the Streamlit frontend and backend, ensuring privacy and domain focus.

Rationale:

- This approach protects user privacy and keeps the agent focused on its intended educational domain.

2. Knowledge Base

Dataset Used:

- The knowledge base is a JSON file (math_kb.json) populated with core math questions and answers.
- Populated via init_kb.py with questions such as:
- "What is the derivative of sin(x)?" → "The derivative of sin(x) is cos(x)."
- "Solve $x^2 + 5x + 6 = 0$" → "The solutions are $x = -2$ and $x = -3$."
- "What is the derivative of log x?" → "The derivative of log(x) is 1/x."
- "Find the area of a circle with radius 5" → "The area is 25π square units."
- "Calculate the integral of x^2 from 0 to 2" → "The integral evaluates to 8/3."

Example Questions to Try:

- What is the derivative of sin(x)?
- Solve $x^2 + 5x + 6 = 0$
- What is the derivative of log x?

3. Web Search Capabilities

How it works:

- If the knowledge base does not have a high-confidence answer, the agent uses the WolframAlpha API for web-based math queries.
- This provides robust, reliable answers for a wide range of math and factual questions.

Example Questions NOT in the KB:

- What is the square root of 2024?
- What is the integral of cos(x)?
- What is the area of a triangle with base 7 and height 3?

Web Extraction Strategy:

- The agent sends the user query to WolframAlpha via API.
- The first result is extracted and returned as the answer, with attribution to WolframAlpha.

4. Human-in-the-Loop Routing for Agentic Workflow

Workflow:

1. Input Guardrails: Check for privacy and domain relevance.
2. Knowledge Base: Try to answer from the local KB using semantic similarity.
3. Web Search (WolframAlpha): If KB fails, query WolframAlpha for a reliable answer.
4. Symbolic Math Fallback: If both fail, use SymPy for symbolic computation (e.g., derivatives, equation solving).
5. Output Guardrails: Check the answer for privacy and domain relevance before di splaying.
6. Human Feedback: Users can provide feedback on answers, which is logged for future improvement.

Rationale:

- Maximizes answer accuracy, privacy, and safety.
- Allows for continuous improvement via human feedback.

- Ensures the agent is robust for both common and novel math questions.

5. Summary Table

- | Requirement | Met? | Notes
- Input/Output Guardrails | Yes | PII and domain checks on both input and output ||
- Knowledge Base | Yes | Populated with core math Q&A, easy to extend ||
- Web-Search-Capabilities |Yes |Uses WolframAlpha API for robust math/factual answers ||
- Human-in-the-Loop Routing | Yes | Multi-step routing, feedback collection, symbolic fallback ||

6. Deployment

- The Math Agent is deployed as a Streamlit web application.
- Users can interact with the agent via a web browser.
- The system is ready for demonstration and further extension.

7. Source Code & Demo

- All source code, configuration files, and setup scripts are included in the project repository.

Prepared by: Lohith Sai Beeram Date: 29-04-2025