

Building & Evaluating Impactful AI Systems

Project overview

Students will design, implement, and evaluate an AI agent system that addresses a real-world challenge in education, healthcare, or environmental sustainability. The project emphasizes building AI agents for genuinely underexplored applications where current AI deployment remains limited, while mandating rigorous fairness evaluation and bias mitigation as core components.

Project Themes and Suggested Applications

Education Domain

Metacognitive learning scaffolding

- Students often lack self-regulation skills; AI tutors typically provide answers rather than developing thinking skills. You'd essentially build an AI system that guides students through planning, monitoring, and evaluating their own learning using methods such as Socratic questioning, progress tracking, reflections etc...
- It should be able to work across different educational backgrounds, learning styles and language proficiencies

Interdisciplinary Concept Connector

- Students struggle to see connections across disciplines; knowledge remains siloed. You have to build an AI system that identifies and explains connections between concepts across different subjects leveraging techniques such as knowledge graph navigation, analogy generation, cross-domain problem mapping, etc...
- It should work across different prior knowledge levels and educational systems.

Healthcare Domain

Health Literacy Translator

- 1 in 3 adults struggle with health information; medical jargon causes errors and poor outcomes. Your goal is to build an agent that translates medical information to appropriate reading levels while maintaining accuracy by incorporating readability adjustment, cultural adaptation, visual explanation generation, etc...
- The system should handle different reading levels, languages, and cultural contexts.

Clinical Trial Matching Agent

Only 3% of eligible patients enroll in clinical trials; matching is manual and inefficient. You have to build an agent that matches patient profiles to relevant clinical trials with explanations by leveraging Eligibility criteria parsing, patient profile analysis, match scoring and ranking while ensuring equitable matching across demographics etc...

Environmental Sustainability Domain

Personal Sustainability Behavior Coach

- People want to be sustainable but lack personalized, actionable guidance. The goal would be to build an agent that provides personalized sustainability recommendations based on individual constraints.
- You can consider features such as carbon footprint tracking, behavior change nudges, alternative suggestions, progress gamification etc...
- Ensure recommendations are feasible across income levels and living situations

Technical Requirements

- LLM Integration with context and memory management, error handling and feedback and prompts with clear templates.
- Web Interface or Command-line Interface
- Identify potential biases and implement at least one bias mitigation strategy.
- Document Limitations and trade-offs.
- (Ideal) multi-Agent architecture with defined roles, clear task decomposition and delegation.

General Tips

- **Start with the problem, not the technology** - Ensure you deeply understand the real-world challenge before coding.
- **Scope aggressively** - Better to do one thing well than many things poorly.
- **Build iteratively** - Get a minimal version working quickly, then enhance.
- **Document as you go** - Don't leave documentation for the end.

Deliverables

- Code repository

- Report in academic format that covers:
 - Architecture, implementation,
 - fairness analysis section,
 - bias analysis with metrics,
 - bias mitigation strategies,
 - clear limitations of current system and recommendations.