

Frontier Models Showdown: Building an LLM Competition Game

Demo and Exploration of Frontier AI Models

- Course focuses on building custom AI tools, not just using existing ones
- Overview of top AI models: GPT-5, Claude, CASA agent, Gemini, Grok, Llama 70B
- Frontier models praised for nuance, synthesis, and reasoning
- Claude highlighted as a community favorite; GPT-5 seen as most powerful currently
- Discussion of trade-offs: intelligence vs. price and speed for business needs
- Links to model resources will be provided

Outsmart Game Demo

- Instructor built "Outsmart" game to showcase LLM competition and strategy
- Four AI players: Alex (GPT OSS 120B), Blake (GPT-5), Charlie (Grok 4), Drew (Claude Sonnet 4.5)
- Each starts with 12 coins; can take/give coins and exchange messages each turn
- Alliances possible—extra coins for ganging up on a player
- Game reveals models' strategies and inner thoughts for transparency
- Game runs slower with larger models, especially GPT-5
- Source code available online; built as a Streamlit app
- Demonstration shows early alliances and evolving strategies
- Alex (open source GPT) takes early lead; later, GPT-5 moves ahead
- Alliances shift: Blake and Charlie team up, Drew (Claude Sonnet) gets targeted
- Emphasis on using model explanations to force deeper reasoning

Key Takeaways and Next Steps

- Exercise provides hands-on feel for model behavior and problem-solving
- Insight into model strengths, reasoning, and interaction styles
- Encouraged to explore models and observe their strategies
- Outcome of the game to be revealed in the next session
- Next session will cover transformer architecture, agentic AI, context engineering, agent loops, and foundational AI concepts
- Participants are 7.5% through the course; encouraged by progress and new understanding of model differences and capabilities