

Testing ChatGPT-5 and Frontier LLMs Through the Web UI

Experimenting with LLM Chat Products

- Focus on direct interaction with LLMs via chat UIs before coding
- Used ChatGPT5 for hands-on exploration
- Emphasis on experiencing structured responses and UI features

Evaluating Business Problems for LLM Solutions

- Asked: How to decide if a business problem fits an LLM solution
- LLM response: define problem clearly, check for ambiguity, assess business value vs. risk, data availability, operational fit
- LLMs strong at structured, open-ended, ambiguous questions
- Provided a balanced, well-organized answer

Comparing LLM Capabilities

- Asked: What questions are LLMs best at, and what's challenging
- ChatGPT5 strengths: teaching, structured explanations, tool-assisted answers, personalization, context memory, synthesis across domains
- Challenges: fresh information, mathematical precision, deep subjective matters, very long reasoning chains
- Recognizes competitors: Claude (long context, human-like), Gemini (real-time multimodal), Mistral (fast, efficient), Cohere (accuracy)
- Acknowledges specialized models may outperform in certain areas

Testing Human-like and Creative Questions

- Explored: "What does it feel like to be jealous?"
- LLM gave nuanced, multi-layered emotional and physical description
- Noted that LLM can draw from training data for such answers
- Encouraged to try more challenging, human-centric questions

Handling Nonsensical or Riddle-like Prompts

- Example: "How many rainbows does it take to jump from Hawaii to 17?"
- LLM responded with a poetic, metaphorical interpretation
- Demonstrated improved handling of illogical or playful questions compared to earlier models

Trick and Meta Questions

- Tested: "How many times does the letter A appear in this sentence?"
- LLM now handles this accurately, whereas older models struggled
- Discussed tokenization and previous model limitations

- Harder meta question: "How many words are there in your answer to this question?"
- LLM provided a correct, self-reflective answer
- Shows significant progress in reasoning and self-analysis

Next Steps

- Will explore more challenging questions and competitor models in upcoming sessions
- Plan to discuss context windows, tokenization, and tool integrations in future lessons