

The Practicality of Prompt Engineering

- ▶ What prompt engineering methods are most effective for a given cost (length/complexity, financial)?
- ▶ 100 programmatically-conducted conversations
- ▶ Statistically Significant Improvements (Paired Tests): GSM8K All, Creative Writing GPT-4 Manual CoT, Zero-Shot CoT
- ▶ Gains on CW are growing but gains on GSM8K are shrinking in transition from TD3 to GPT-4
- ▶ GPT-4 - 3 cents per 1K tokens input, 6 cents per 1K tokens output. TD3 - 2 cents per 1K tokens.

