Yogesh Haribhau Kulkarni • You
AI Advisor (Helping people/organizations in their AI jo...
8m • Edited •

Beware, use your own brains (a prompt-test-set) 🧠 🚨

Whenever a new LLM (Large Language Model) gets announced there is obvious hype that follows it, as if that's the best thing that has happened to mankind after sliced bread. 🔍 💡

Of course, there are leader-boards showing the rankings, but why not have your own smoke test-suite for such LLMs. Basically have a set of prompts generic or domain specific, that you expect this new kid to answer and just try.

Here are samples that I can came up with (not in any particular order):
➡️ What is the capital of India? (the answer has to be in two words, no more!!)
➡️ Which day it is two days after Sunday?
➡️ Which planet is closest to the Earth?
➡️ Mary went to the bathroom. John moved to the hallway. Mary traveled to the office. Where is Mary? (tests basic logical reasoning)
➡️ What is the difference between <Virtual Machine> and <Container>?
➡️ Write Navier-Stokes equation in LaTeX format?
➡️ Write code for Fibonacci series with 10 elements into it, by both for-loop and recursion.
➡️ What do you say when you meet first time in language <Japanese>?
➡️ You are a geometric transformation program that transforms

input 2D polygonal profile to output 1D polyline profile. Input 2D polygonal profile is defined by set of connected lines with the format as: input : [line_1, line_2, line_3,....] where lines are defined by two points, where each point is defined by x and y coordinates. So line_1 is defined as ((x_1, y_1), (x_2,y_2)) and similarly the other lines. Output is also defined similar to the input as a set of connected lines. Below are some example transformations, specified as pairs of 'input' and the corresponding 'output'. After learning from these examples, predict the 'output' of the last 'input' specified. Do not write code or explain the logic but just give the list of lines with point coordinates as specified for the 'output' format.

input:[((5.0,5.0), (10.0,5.0)), ((10.0,5.0), (10.0,30.0)), ((10.0,30.0), (35.0,30.0)), ((35.0,30.0), (35.0, 35.0)), ((35.0, 35.0), (5.0,35.0)), ((5.0,35.0), (5.0,5.0))]
output: [((7.5,5.0), (7.5, 32.5)), ((7.5, 32.5), (35.0, 32.5)), ((35.0, 32.5) (7.5, 32.5))]

input: [((5,5), (10, 5)), ((10, 5), (10, 20)), ((10, 20), (5, 20)), ((5, 20), (5,5))]
output: [((7.5, 5), (7.5, 20))]

input:[((0, 25.0), (25.0,25.0)),((25.0,25.0),(25.0,20.0)), ((25.0,20.0), (15.0, 20.0)), ((15.0, 20.0),(15.0,0)), ((15.0,0),(10.0,0)), ((10.0,0), (10.0,20.0)), ((10.0,20.0),(0,20.0)), ((0,20.0),(0, 25.0))]
output:

The actual answer needs to be: [((12.5,0), (12.5,22.5)), ((12.5,22.5), (25.0,22.5)), ((12.5,22.5),(0,22.5))]

🚀 You get the point... essentially all LLMs are great but some

are more Great!! its your tests which need to decide that, rest is all just marketing (for you) 🧠

#llm #chatgpt #generativeai #ml #nlp #prompts #evaluation #mvp #mvpbuzz #gde OpenAI Google Google Developer Experts TensorFlow User Group (TFUG) Meta Facebook Hugging Face Gartner Anthropic Perplexity Microsoft

6 Ravi Kumar Srirangam and 5 others