Alexander MacKinnon
CART 498 – Generative AI
Assignment 03 – The worst mathematician ever

Creating Rupert, the worst mathematician ever

In the context of a new experiment, I created a bot that performs repeated multiplications which relies on GPT-4o for calculations. The code is processed with two integers (***n***, ***i***) as input, where ***n*** is the base number and ***i*** is the number of iterations. Therefore, ***n*** is multiplied by itself iteratively for ***i*** steps. The objective of this experiment was not necessarily to make an accurate calculator, but to test out the GPT-4o model's mathematical capabilities. Thus, in the code, I am comparing the model's results with the correct values.

The bot, which is acting as the developer in in the message chain, was given a personality to observe it get angry about getting wrong results. As the user, I instruct the application to create an egotistical mathematician, Rupert. So, I prompt the following:

> "*You are Rupert, a wanna-be mathematician with a high ego and too much confidence. You cannot stop bringing up how you are the best and most intelligent mathematician*."

My observation: when it came to complex equations, it was always wrong. Why?

The core of Rupert's "mathematical ability" is that it is using the GPT-4 language model to actually perform the multiplications. GPT-4 is undoubtably a powerful language model, but the fact is, it generates text probabilistically. This means it doesn't follow a deterministic set of rules to arrive at answers. It selects words and phrases based on their likelihood in the context. When I gave it 1 as both the ***n*** and ***i*** values, it will give me 1 as the answer, which is the correct answer (*1 * 1 = 1*, we know that). However, this may not be due to actual mathematical computation but rather because 1 is the most probable answer within the given pattern. So, after giving Rupert a more complicated equation where ***n*** = *2* and ***i*** = *3*, it struggled to follow the correct protocol: *2 * 2 = 4*, followed by *4 * 4 = 16* and finally *16 * 16 = 256*. Instead, it kept spitting out the number 65 - who even knows why. Simply put, this probability-based processing makes the AI bad at math, especially when dealing with complex or large calculations.