# Assignment 3 Reflection

Leah Song
40321005

Prompt: A short reflection (250-350 words) explaining why your bot is failing. If it fails at all! Submit it as a PDF file.

I had a lot of fun working on this assignment. At first, I struggled to understand how the purpose of this task connected to what we had learned previously. However, as I progressed through the assignment and began to understand how large language models (LLMs) actually work. It helped clarify many of the concepts we discussed in class. Experimenting with different parameters such as temperature, top_p, input structure, and model choice help me understand better how these variables influence AI behavior.

Before this assignment, I encountered content claiming that AI is bad at math. Conceptually, this did not make much sense to me at first. I assumed that since AI receives a prompt describing a mathematical operation, it should be able to give the correct answer. Through this assignment, however, I learned that LLMs are in fact not reliable for simple or mathematical operations. A traditional programming language, such as Python, is more suitable for this type of computation.

Models like GPT are designed primarily for language generation, not numerical computation. The model does not actually execute mathematical operations. Instead, it predicts what output is most likely to follow based on patterns learned during training. In my implementation, the model was also running with a high temperature value, which increased randomness in the output. A lower temperature would have been more appropriate for tasks that require precision.

Additionally, the model was instructed to perform multiple calculation steps while only outputting the final result. Since none of the intermediate steps were verified, even a small error could lead to a significant discrepancy in the answer. This issue is emphasized by the fact that the mathematical expression in this assignment grows exponentially, even with small input values. When numbers become too large, the model may struggle due to token limitations or simply lose accuracy when handling long numerical strings.

Overall, this assignment helped me understand that AI is best used for interpretation, explanation, and interaction rather than precise arithmetic. While the sarcastic, self-deprecating comments were hardcoded in my program, I would have liked to explore ways to have the AI generate its own personality-driven responses.