Kamyar Karimi
CART-498
Gabriel Vigliensoni
2025-02-05

On Mathematical Precision and Transformers

Given the nature of transformers in machine learning, they are primarily designed and useful for language processing and predictive tasks. These models are designed to predict the next word in a sequence by leveraging multi-dimensional matrices and complex attention mechanisms, and given their statistical essence, they're not the best and most optimized agents for mathematical or arithmetical tasks. Given that there's no genuine 'understanding' of mathematics in these models, they mimic (as transformers do in all of their generations). Hence, they fail at mathematical operations.

In this assignment, we demand the model to iterate through multiple mathematical operations and the notion of "tracking" numbers, that I did in the python script (using the for loop for example), by the model is almost non-existent as the exponential computation becomes harder to follow. Especially the fact that transformers always predict estimations and approximations makes them less reliable for mathematical tasks as they're not designed to be 100% logically precise, which is the foundational requirement in arithmetic mathematics. This probabilistic approach contrasts sharply with the strict logical precision required in arithmetic operations.

The inherent limitations of these models mean that while they can generate language that appears mathematically reasonable, they are prone to errors when actual numerical precision is required. In this case, as the operations become exponentially more complex, the model's approximations diverge from the exact results produced by an algorithm. Therefore, even if the prompt includes detailed instructions for iterative multiplication, the model's response will likely be an estimation rather than an exact answer, highlighting the inability of transformer-based models for tasks that require strict numerical precision.