In-Context Learning as Optimization

Topic

The Meta-Learning Dynamics of Large Language Models

Background

Large Language Models demonstrate a remarkable ability to adapt to new tasks via Prompt Engineering and In-Context Learning (ICL) without undergoing explicit parameter updates.

Assignment Task

Provide a formal justification for framing In-Context Learning and Prompt Engineering as a form of implicit gradient descent. Explain the theoretical mechanisms by which the forward pass mechanisms, particularly self-attention over the context window, mathematically or conceptually emulate the parameter updates typically achieved through backpropagation.

Submission Expectation

Prepare a rigorous, self-contained written response that defines all key assumptions, uses precise technical terminology, and supports the argument with mathematical, architectural, or conceptual reasoning where appropriate.