Cover Letter
You Didn't Train a Model: Common Patterns in Mislabeling AI Development Efforts

Anonymous Author(s)
Affiliation
Address
email

1 Expected Watch Time

6 minutes and 24 seconds

2 Target Audience

Those who make decisions around AI tool integration without being themselves researchers or experts. This includes both buy-side and sell-side practitioners such as industry product leads; potential product implementers; policy, safety, and governance practitioners; educators who teach LLM fundamentals.

3 Brief description of the material

I am pleased to submit the video *You Didn't Train a Model: Common Patterns in Mislabeling AI Development Efforts* for consideration. This short piece is intended to clarify a pervasive misunderstanding in AI discourse: the difference between training a model and using an existing model through prompting, retrieval, orchestration, or fine-tuning. While the distinction may seem semantic, it carries practical implications for reproducibility, governance, and public trust.

The material begins by defining what model training entails, grounding the concept in accessible language while accurately describing the process of starting from randomly initialized weights and refining them through repeated prediction, evaluation, and adjustment. It illustrates the scale of true training with concrete examples, such as Meta's Llama-3.1 at 405B parameters and GPT-4's reported nine-figure compute cost. These examples highlight why genuine training is a resource-intensive endeavor accessible to only a small number of organizations.

The video then contrasts this with common and valuable engineering practices that do not involve training base model weights: prompt design, dynamic context injection, retrieval-augmented generation, multi-model routing, and parameter-efficient adapters. While these approaches can yield substantial product impact, representing them as "training a model" can mislead stakeholders about intellectual property, technical capability, and associated risks.

Finally, the video explains why precise terminology matters: it enables clearer reporting, ensures reproducibility, sets accurate expectations for investors and policymakers, and supports responsible deployment in sensitive domains. It concludes with practical, accurate language teams can adopt and a concise litmus test for distinguishing training from usage.

I believe this work aligns with NeurIPS's call to create educational materials that introduce AI concepts (training) to non-expert audiences. Thank you for your consideration.