

Memorandum on LLaMA2 Fine-Tune for AI Startup

Introduction

This is a memorandum on the fine-tuning of LLaMA2 for the purpose of creating a competitive AI startup. The goal is to gather deep technical insights and prescient intuition on the nature of fine-tuning in relation to growing a flourishing business in the age of AI. Analysis will be provided on the principles of fine-tuned LLMs and their application to the firm.

It is my belief that an important memo must encapsulate the essence and importance of an subject matter. It must further provide a framework for understanding the subject matter and it's implications. It should be a document that is remembered and referenced. The document should suggest actionable plans for the future. It should be forward thinking and forward looking.

To this effect, I leverage the book, *"Competing in the Age of AI"*, which provides many case studies in the nature of AI firms; it summarizes the economic landscape of AI, teaches the basic fundamentals of business, and additionally offers many a predictions for the course of AI. The business model and operating model of the LLaMA2 fine-tune will be contextualized in the framework set forth in the book. I will be providing a brief analysis in the practical, actionable, and build-able applications that can arise from a LLaMA2 fine-tune, to the purpose of formulating a workable business and operational model.

Technical Details - Insight and Intuition

The principles laid forth by this memo is synthesized from the experience gained through the implementation of a LLaMA2 fine-tune. The technical details of the fine-tune are as follows:

- Training data is sourced from Kaggle and synthetically populated with OpenAI'S ChatGTP-4
- Fine-tuning is then executed on Colab's compute using code available from the Open Source community, particularly Github
- The resultant fine-tuned models are then uploaded to HuggingFace, which allows for leveraging it's APIs for inference and evaluation
- Inference is then used to generate the new fine-tuned data
- Evaluation is performed through eye-tests and a haphazard benchmark

This implementation stands on the shoulders of existing AI/ML giants and their ecosystems: e.g. Kaggle for datasets, Colab for training, Meta & Hugging Face for models, Github for repository, and Replit for presentation, etc.

However, this particular process and pipeline is self-devised and manually implemented respectively. This exercise in integrating various tools and techniques leaves the impression of very possible and potential value add - Business Scale can be achieve through the automation of manual pain-points, through the application of applied software engineering. Increasing Business Scope will require an concerted effort to understand customer data on a case-by-case basics, but that too can be formalized, leading to further automation. Lastly, Business Learning is achieve through an attachment and mutualism with the AI community at large, possibly through formal partnerships and collaborations.

Proposal - What We Can Build

LLaMA2 fine-tune is a stepping stone for better and more advanced AI systems. I believe fine-tuning is but a very narrow use case, and one of many methods, to extract useful business use cases from existing foundational models. However, I also believe all these techniques build towards ever more performant systems. Recent development in multi-model technologies substantiate my belief that the world of tomorrow will be built on the foundational models of today.

To this end, I propose developing LLaMA interns - AI systems fine-tuned on our data to execute specialized tasks. Just as with human interns, These LLaMA interns would acquire domain expertise through training. With the right data and methods, LLaMA interns could perform knowledge work; in building business processes around LLaMA interns, we can automate and scale our operations as we grow - casting us as an AI first company prepared to handle the rapidly changing business landscape.

As when onboarding employees, imbuing domain knowledge in LLaMA interns requires patience - they must be guided gently to grow into their role. However, once trained, LLaMA interns can be integrated throughout operations much like humans are trained in operational processes. Most importantly, the clonable nature of LLaMA interns will truly allow us to scale our operations as we grow. Further assessment into the underlying business process, with an focus on automation, will inevitably confer advantage for the firm.

My LLaMA2 fine-tune is but a tongue-and-cheek trail run in building such a system, but underlying the experiment is an idea that I believe is worth pursuing.

Role of AI-Human Collaboration

We need to build a firm that maximizes the collaboration between human and AI intelligence. Within interactions is where the spark is... it's where ideas are generated, where inspiration comes, and where the human consciousness expands. If we can build such environment, we have a competitive advantage in the age of AI.

We will need an ethical approach between keeping humans in the loop and removing humans from the loop: balance the need to ensure the values of the company is upheld and the need to gain a competitive advantage through AI automation. Reinforcing our companies mission statement amongst ourselves and amongst the AI systems will fortify us during the trying times.

We should be present and mindful - while bringing AI technologies to market. Attention is what we'll need ... to model patterns of AI systems as they emerge and be responsive to business needs. We risk missing key insight from these AI systems should we be lax. Additionally we should frame our work around enjoying the creative process of building and experimenting with AI systems. We should enjoy our time playing with state-of-the-art technologies ever created and not just aim for uncertain results.

We will need to find community. We will need to find partners. We will need to find collaborators. We will need to find customers. We will need to find employees. We will need to find investors. We will need to find mentors. We will need to find friends. We will need to find family. We will need to find love. We will need to find ourselves.

The Human Element

As the firm expands, it is crucial we remain thoughtful about the human element. There is understandable hype and anxiety about how these technologies will impact society. We must balance the promised benefits with potential disruptions to jobs, businesses, and social structures. In the face of the disruptions, I would affirm that AI should ultimately help us expand the range of positive human experience ... enhance our ability to enjoy life, love, and work: humor, humility, fun, excitement, mental stimulation, being aware, being present, being mindful. Questions we must ask are: How does AGI help create a better world? More importantly how does AGI help spread the vibes of a better world?

Nature of the Company

I will advocate for building a different kind of company.

On the business front, we should possess and secure the ability to abstract our underlying business processes. We should focus on digitalizing our business processes to better enable AI and automation. In that, we will have an advantage in analysis and deploying AI solutions as economically as possible in a deliberate step by step fashion. On the technical front, we can scale faster, innovate faster, and personalize better by applying AI, analytics, networks, software, and automation. We will need to develop and procure sophisticated methods of gathering and using data. We'll need fast ways to react to information, better ways to make informed operating decisions, and efficient ways to execute tasks.

In building the firm we must constantly ask: "How does our company flourish in the age of AI?". A focus on deliberate use of AI in our standard operating procedure should be forefront when we architect our firm so that we can shape the right skills, capabilities, and culture to drive an increasingly digital/AI-centric operating model. We must weave AI into the fabric of our firm. We need to build a company that has an AI/ML foundation.

Forward Thinking - Analysis

As a startup hoping to bring AI technologies to market, we need to stay on, or be adjacent to, the leading edge of AI innovation. It'll be necessary for us to set trends that propel the firm ahead, while creating tools that accelerate human progress. I believe we have the capabilities to make a valuable product and be rewarded for it. We as individuals are banding together as a company to accomplish a goal that we cannot accomplish alone. It is my belief that building together as friends who have the shared goals and as professional peers with complementary skills will surely lead to success.

The business landscape is shifting, but by embracing AI, we can thrive in this new era. Our future will be defined by how we steer AI's potential today. It's an exciting future for us to navigate.