Playing around with LLMs

HSF-India Scientific Computing workshop: IIT Madras Jan 5-9, 2026

Slido **(Feel free to ask questions here, it's anonymous, so no need to feel nervous)**

Code: 1170874

Link active till Friday, so you can ask questions over the next 3 days, goal is to learn and have fun :)

What are LLMs?!

- Large = Billions of parameters (in yesterday's deep learning class we trained a model with O(1000) parameters)*

- Language = Inputs to model are human language

- *Nowadays, models with around 7 billion are also called small…

- Eg. ChatGPT's, Gemini's, Sarvam's backend models

Overview

- LLMs trained on petabytes of data (many of the models we use don't openly state what data they use, including some open-weight models)

- They generate the next word (or part of a word called 'token') based on a proabability distribution over the previous words in a sentence.

- They are memory machines

- Have Jagged intelligence: So make sure to never trust them blindly

- Some of them trained over several months on a large number of GPUs (graphical processing units)!

How do we 'teach' them language

- The underlying magic is Attention

- Modern LLMs use an architecture called transformers (which we saw in the ML lectures)

- Eg. In the sentence "Z' is a hypothetical boson like the known Higgs Boson, it has not yet been seen experimentally". LLMs correctly predict that 'it' refers to Z' and not Higgs.

- Typically data is scraped from the internet*

- *along with all the biases, false info etc…

LLMs into assistants

- At the end of 'Next word prediction' training (pretraining), it has learnt the structure, grammar etc. But it's not good at following instructions

- Further training is done using question answer pairs designed by humans (supervised fine-tuning)

- And then more training is done using thumbs-up, thumbs-down kind of stuff to match human preferences (reinforcement learning)

Things to keep in mind

- Capabilities: Translation (depending on the resource available for training), summarization, coding, story telling

- Trap: Model is probabilistic, not factual !

- Hallucination: Confidently states something false, because it 'sounds' probable.

All this is good, what's in it for physicists?

- Assumption: Physicists are normal people (maybe?)

- LLMs are a recent, really powerful tool

- People are figuring out how to use them

- Goal for today: Get comfortable using LLMs, so that you are comfortable with this tool. Next time when you see some problem, you won't feel nervous if you want to investigate using LLMs to solve it.

- One example: We can use LLMs to build systems which help make collaborations more efficient : Demo

Hackathon
Prerequisites-1

- What is the rank of a matrix? Is the rank of a 3x3 matrix always 3?

- How does sampling work?

- What does supervised training mean?

- What is dropout? Why is it used?

- Python virtual environment?

- Is the dataset size important for quality of training?

- Can I just copy a single data entry 500 times and say I have 500 training events? Why or why not?

Hackathon
Prerequisites-2

- Create an free account on [HuggingFace](); Settings->Access tokens-> Create new token, give it a name you like.

- Use the read option (we don't need write as we are just using existing models)

- Don't share your token with anyone!!

- Copy the token once you create it (the long string, not the token name).

- Paste this in the first cell of Jupyter hub in place of "Your_HF_Token".

Hackathon: Use this [link]

Task and things to think about

- Can we tweak a LLM to make it speak in a certain style?

- Eg. One of my friend uses the word "acha" often. Can we make the LLM use the word "acha" in every answer it gives?

- Use generate.py to create the dataset that we use for training

- Then run the llm_sft.ipynb notebook

- Things to think about:

  ✓ Do you see any side effects? How can we fix it?

  ✓ What happens if we skip the model.eval() cell? Why?

  ✓ Anything else you found that you want to share with all of us?

thank you