Machine Learning for the Skeptical HEP Theorist

Jesse Thaler (MIT | IAIFI)

Brown HEP Seminar

Nov 7, 2022

Great to be back in my old stomping grounds!

Many a happy memory in this room!

We've all seen the power of AI

"chalkboard physics equation" $\Longrightarrow$ $\boxed{E=mc^2}$

But can we use it for scientific applications?

AI is not magic!

At some level, AI is "just" statistics with strong $\underbrace{\text{inductive biases}}$.

↑ not a familiar term for most of us

For DALL-E:

$$p(\text{images}, \text{descriptions})$$

all possible sentences

all possible images

all possible mappings

Via Bayes theorem:

$$\underbrace{p(I \mid D)}_{\text{what DALL-E estimates}}\; p(D) = p(D \mid I)\; p(I)$$

learned from data and/or baked into function

E.g.

gain an "understanding" of this operation

"Reversed chalkboard physics equation" $\Longrightarrow$ $mc^2 = E$

This is a hard problem, but not insurmountable!

For physics, we often have strong inductive biases regarding symmetries

$$p(\text{jet} \mid \text{"quark"}) = p(\text{reshuffled jet} \mid \text{"quark"})$$

permutation symmetry.

$$\text{jet} \Rightarrow \{ p_1^\mu, p_2^\mu, \cdots p_N^\mu \}$$

"Easy" to make this permutation invariant

$$p(\text{jet} \mid \text{"quark"}) = F\left(\sum_{i \in \text{jet}} \Phi(p_i^\mu) \right)$$

$\Phi : \mathbb{R}^4 \to \mathbb{R}^L$ ← latent space

$F : \mathbb{R}^L \to \mathbb{R}$

You can show that this is complete for sufficiently flexible F and Φ.

Requirement for AI:

- Well-specified problem ← rest of this talk
- Reliable training data
- Learnable function (e.g. NN)
- Powerful optimizer ← related to quantum many-body physics!

Physics input necessary for all of these!

But my "phase transition" was about finding a problem specifications, so I want to focus on that.

Key insight ...

Lagrangian mechanics! (Really?!)

First, let me remind you of Monte Carlo integration.

$$\int_0^L dx\, f(x) \approx \underbrace{\frac{1}{N}\sum_{i=1}^{N}}_{\text{uniformly sampled in } x \in [0, L]} L \cdot f(x_i)$$

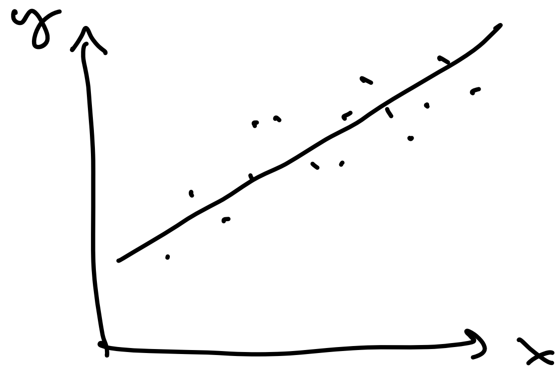
For non-trivial probabilities:

$$\underbrace{\frac{1}{N}\sum_{i=1}^{N}}_{x_i \text{ sampled from } p(x)} f(x) \approx \underbrace{\int dx\, p(x)}\ f(x)$$

Physicists love integrals! $\Rightarrow$ Lagrangians

We tend to not love sums, but for "enough" samples they are the same.

In this language, what is linear regression? 6



$f(x) = mx + b$

$$\mathcal{L}[f] = \int dx\, dy\, p(x,y)\, \big(y - f(x)\big)^2$$

Want: $\underset{m,b}{\operatorname{argmin}}\ \mathcal{L}[f]$

For HEP applications, we often want to do hypothesis tests:

$$p(\{x\} \mid A) \quad \text{vs.} \quad p(\{x\} \mid B)$$

Because collider data is "IID"...
(independent and identically distributed)

$$p(\{x_1, x_2, \ldots\} \mid A)$$
$$= p(x_1 \mid A) \cdot p(x_2 \mid A) \cdots$$

So all we need for hypothesis testing is

$$\frac{p(x \mid A)}{p(x \mid B)} \leftarrow \text{Neyman - Pearson Lemma}$$

Switching notation to make it easier to read:

$$\frac{p(x)}{q(x)}$$

Given samples P and Q, can I learn a function $f(x)$ that approximates likelihood ratio?

If you already know $p(x)/q(x)$, no need for AI!

But if you only have samples from P and Q:

$$\mathcal{L}[f] = \int dx\, p(x) \boxed{A[f]} + \int dx\, q(x) \boxed{B[f]}$$

such that "Euler-Lagrange equation"

$$\frac{\delta \mathcal{L}}{\delta f} = 0 \implies f(x) = \frac{p(x)}{q(x)}$$

Many answers!

$$\frac{\delta \mathcal{L}}{\delta f} = 0 = p(x)\, A'[f] + q(x)\, B'[f]$$

$$\frac{p(x)}{q(x)} = -\frac{B'[f]}{A'[f]}$$

e.g. $A[f] = \log f(x)$

$B[f] = 1 - f(x)$

Going back to stats notation:

$$\mathcal{L} = \langle \log f(x) \rangle_p + \langle 1 - f(x) \rangle_q$$

Asymptotically:

$$\underset{f(x)}{\arg\max}\ \mathcal{L} = \frac{p(x)}{q(x)}$$ ← Likelihood ratio

$$\max \mathcal{L} = \int dx\, p(x) \log \frac{p(x)}{q(x)}$$

Kullback-Leibler divergence.

Many other choices corresponding to different approaches in ML literature.

If you can write a loss, you can leverage AI/ML!

Interpretability, uncertainties, symmetries, ...
Lagrange multipliers, non-locality, ...

So why is AI so mysterious?

Because recent advances convolve...

1. problem to be solved with
2. algorithm to find solution.

Physicists can contribute to both!

But the key (for me) is to consider them separately:

1. ← HEP inference
2. ← Quantum Many-Body Physics

Ask me lots of questions!