

Hello 2026

Poojan Shah

As is common with research, I have no idea where 2026 will take me, but I do know where I am going to start from. I'm writing down some high level problems I want to think about and see if I can (atleast partially) answer them by this year :

1. The success of deep learning is often attributed to the so-called manifold hypothesis, which posits that high dimensional data concentrate near a low dimensional manifold, or a union thereof. Can this structure be exploited to design faster algorithms ? How can this structural assumption on data be used for *beyond worst case algorithm analysis* ?
2. How to define the intrinsic dimension of data ? How to measure it ? Is it statistically feasible ? computationally tractable ? Is there a stat-to-comp gap ? What amount of noise destroys low dimensional structure ? Can modern algorithmic techniques allow us to explore these questions ? What is its relation to deep learning ?
3. Why does self-supervised / multi-task learning work ? Can we quantify the relation between the proxy objective and true objective ? What is the power and role of *shared representations* ?
4. For what kind of tasks and problems do direct product theorems exist in learning theory ?

If you want to share any ideas or collaborate in trying to answer these questions - please reach out to me on my mail !