

UNet Reflection

GANs were a lot harder to train than our diffusion model, due to their instability (my model kept going into mode collapse). This is due to the minimax game involved in GAN training. On the other hand, however, our diffusion model was a lot more computationally intensive than our GAN. It was also a lot harder to implement (more lines of code, needed a stronger understanding of what I was actually doing, etc., although this could in part just be due to me understanding GANs better). In terms of our results, in the cases that our GAN did not fail in training, it produced better samples than my diffusion model did.

Towards Robust Out-of-Distribution Generalization Bounds via Sharpness

Summary: This paper provides generalization bounds for OOD generalization based on the sharpness of the loss function minimum. The authors produce this robust generalization bound through partitioning the sample space and measuring the distribution shifts, and produce the theory that a minimization of this sharpness can lead to better OOD generalization.

Strengths/Weaknesses:

- Very easy to follow. A lot of math that is seemingly rigorous at first, but very well explained. Math also makes logical sense – even though the proofs were a bit questionable at some points, it generally followed.
- I vaguely recall reading some paper in which it actually suggested that sharper minima was better for OOD – perhaps extending their work to more datasets outside of RotatedMNIST and whatever synthetic dataset they had would be nice
- Logical idea, sharp min bad, flatter min good and better for OOD generalization
- I feel like I'm missing the big picture of figure 4. What is the key insight we are presenting here?

Clarity: Very clear paper. Very straightforward, digestible, and easy to understand. Experiments can be understood well, seems to be very thorough.

Quality, Novelty, Reproducibility: Quality could be improved, re: extending to more datasets (it also seems a bit like the conditions are explicitly chosen, I wonder what results would look like changing it up a bit). Pretty novel paper, it introduces a somewhat

new metric for OOD generalization bounds. I think I could reproduce these results, given a better understanding of the math behind it.

Recommendation: 7. Looks solid to me.

Optimal Sample Complexity for Average Reward Markov Decision Processes

Summary: This paper develops an estimator for the optimal policy of uniformly ergodic average reward MDPs (AMDPs) with some sample complexity, under a generative model.

Strengths/Weaknesses:

- The result of closing the gap between the upper and lower bound of the sample complexity for AMDPs to learn an optimal policy is (to me, a stranger to the field) pretty important.
- A little bit weak, the reduction from AMDP to DMDP seems to have existed for a while, and that this result had just been something that had mainly been glossed over.
- The approach is simple, and makes a lot of sense (however if this reduction has existed for a while like I said in 2nd bullet, it better make sense).
- It is unclear to me how important this result for the wider field of RL when this applies specifically to uniformly ergodic AMDPs.

Clarity: Seems pretty clear. I could mostly follow, even though I am not familiar with the field.

Quality, Novelty, Reproducibility: Was an okay paper. Everything made sense, I could follow along well. Novelty didn't really seem to be there, see comments in weaknesses. Reproducibility is low, I would NOT know how to turn this into code.

Recommendation: 6. Good enough to pass, just doesn't seem super amazing.