CS 4644/7643: Deep Learning
Spring 2024
HW2 Solutions

Arthur Scaquetti do Nascimento

Due: 11:59pm, Feb 19, 2024

7.a)

- Key contributions: This work clearly lays out a well-defined framework for gauging the effective capacity of machine learning models. Even though the authors do not address the shortcomings of Deep Learning that they found, they do state that engineering a (deep) network to perform optimization is not a hard task despite of the models not being able to generalize. Moreover, the authors state that this is an indication that the very reasons that make optimization easy in practice must not be why generalization works.

- Strengths:

  - The design of an evaluation framework for "measuring" the effective capacity of ML models (i.e., looking at randomization tests, explicit and implicit regularization);
  - The very questioning tone, leading to 'how can we trust that blackbox models are generalizing other than looking at test and validation curves?'. That point is very important, because it brings to light that sometimes we are looking at the wrong metrics, even though those metrics have been extensively used in the field.
  - Bringing to light implicit regularization, which is often overseen, and was particularly unknown to me.

- Weaknesses: The main thing that unsettles me about this type of paper is that they do not propose an answer, or designing guidelines such as a structured methodology in which it is guaranteed to only optimize when generalization happens. I know, this is basically solving the whole blackbox problem at once, but it bugs me when a work is published "just" to point out flaws. However, I think this work would have been a dozen times stronger if it had some design guidelines (even if faint), or proposed a different metric to look at.

7.b) Personal takeaways:

I had never thought of the impacts of implicit regularization, and will definitely keep an eye out for that. I also had never considered that Deep Learning models could present such good performances using completely corrupted datasets – this is potentially one of the things I will carry the most from reading this work: when in doubt if a model is actually good, I will try it out on a noisy and/or randomly labeled dataset as a sanity check.

However, other than that, I do not think that I gained that much insight from this work. Unfortunately, we still need to rely on blackboxes for the sake of solving certain problems. Hopefully XAI helps us with that.