Too Salty: On Model Souping and Neural Averaging

Pierre Mackenzie * **Simon Ghyselincks** *

Abstract

Souping Abstract

1. Introduction

1.1. Motivation

Why do we care about souping? Wortsman et al. (2022)

1.2. Related Work

Other works on souping and neural averaging. What is/isn't understood?

2. A Theory of Souping

2.1. Why does Souping Work?

Theoretical motivation. Averaging out of noise. Does this lead to a prediction of how much gain in performance we should expect?

2.2. A Toy Model

Demonstrate noise reduction on a theoretical model (high-dim Guassian?), and a toy model.

2.3. Defining Soupability

What does it mean for models to be soupable? Define metrics for both in-distribution and out-of-distribution soupability.

3. Experiments

3.1. Method

What models did we use? What data? What loss function? On what cluster/GPU? How long did it take to run? How many models?

3.2. Predicting Soupability

What experiments did we run to test our ability to predict soupability? How well did we do? What correlates/doesn't? Is there stuff that's cool to put in the appendix?

What about predicting soupability of A to C given the soupability of A to B and B to C?

3.3. Validating the Theory

What experiments did we run to validate our theory of souping? Do we get the noise reduction we expect?

4. Conclusion

What did we find out? What did we learn? What are future avenues of research that might help?

References

Wortsman, M., Ilharco, G., Gadre, S. Y., Roelofs, R., Gontijo-Lopes, R., Morcos, A. S., Namkoong, H., Farhadi, A., Carmon, Y., Kornblith, S., and Schmidt, L. Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time, 2022. URL https://arxiv.org/abs/2203.05482.

A. Appendix

Anything else?