DATA7703 - Machine Learning
Case Study 2020
(worth 10% of total course marks)

Marcus Gallagher

May 2020

1 Introduction

In the course you have heard about the core trade-off that exists in machine learning between under/over fitting, model complexity and training set size. A recently-published paper[1] suggests that this is not an adequate explanation of over-parameterized models, including deep neural networks. Their idea is summarized in Figure 1.

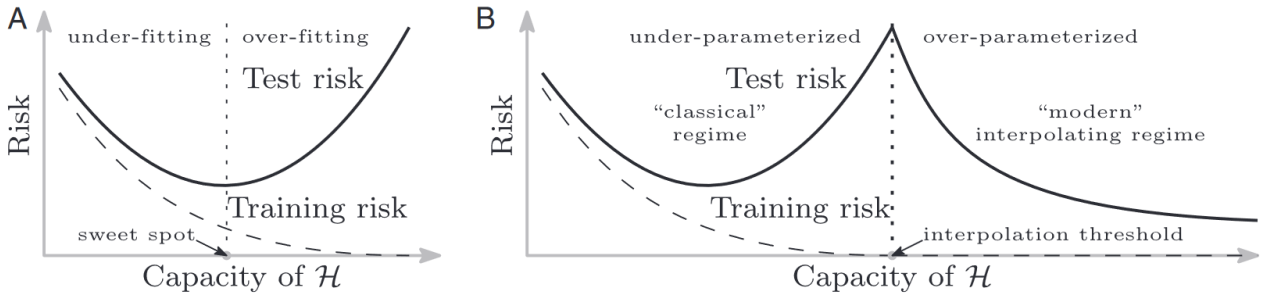

Fig. 1. Curves for training risk (dashed line) and test risk (solid line). (*A*) The classical U-shaped risk curve arising from the bias–variance trade-off. (*B*) The double-descent risk curve, which incorporates the U-shaped risk curve (i.e., the "classical" regime) together with the observed behavior from using high-capacity function classes (i.e., the "modern" interpolating regime), separated by the interpolation threshold. The predictors to the right of the interpolation threshold have zero training risk.

Figure 1: Figure 1 from the paper!

Your task in this case study is to design and carry out a scientific experiment that tests the phenomena of double-descent curves in machine learning. This is an individual assessment task, so your submission must be your own individual work.

2 Procedure

- Read the paper carefully.
- Design an experiment/set of experiments.
  - You will need to decide what kind of model(s) and dataset(s) to use, as well as make decisions about the hyperparameters of the models and parameters of your experiments (e.g. number of trials to perform).
  - You are free to use any code you have written during the course, software libraries and software packages to implement your experiments, as long as you appropriately cite the work of others. You probably shouldn't spend time implementing models/algorithms from scratch to do your experiments with - marks are based on the scientific experimental design, the results and your write-up of the experiment. However, you definitely should come up with your own experimental design from scratch. Taking (for example) a

blog post where someone has done something different and tweaking it to fit this task is unlikely to receive high marks.

  - Experiments can take a significant amount of computation time. You are not expected to do anything that requires more than (for example) 24hrs of compute time on a typical desktop or laptop computer.

- Carry out your experiments, collect and analyse the results.

3 To Hand-in

Write up a brief report (at most 5 pages of text, not including figures and tables) describing what you have done and the results.

Note: this is a scientific experiment - I do not know if you are going to be able to successfully replicate the double-descent curve that the paper hypothesizes. It is likely to depend on many experimental factors. But you should not stress about this - achieving a high mark is not about whether your experiment is successful! Rather, it is about the quality, rigour and presentation of your work. I am interested in reading about what you did, why you did it, what can be concluded for your experiments and what you learned.

4 References

[1] Belkin, M., Hsu, D., Ma, S., Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. Proceedings of the National Academy of Sciences, 116(32), 15849-15854.