Final projects - EE-411 - 2024-2025

December 16, 2024

Abstract

This document contains the list of the final projects for the course "Fundamentals of Inference and Learning" for the Academic Year 2024/2025.

1 Introduction

The final project for this course will consist in a **Reproducibility Challenge**. That is, you will have to *critically reproduce the results from a published machine learning paper*. To be precise, we expect you to reproduce a couple of plots or tables from an established paper. Overall, the tasks for your final project assignment are:

1. Attempt to reproduce numerically the results presented in the paper using the same settings described in this work, or slightly different ones.
2. Explain why these results are important and put them into context.
3. Review critically your attempt at reproduction, and discuss critically the paper in view of your own results.
4. Propose new research directions based on your understanding.

2 Project guidelines

- We wish to give as much time as possible for this project, so you have until January 31st to send your result. However, we strongly advise you to finish it (much) sooner not to interfere with your exams! This project will account for 40% of the final grade. There are a limited number of project to choose from, and you will not be able to choose other projects. You should use the project as an opportunity to "learn by doing".
- You may work in **teams of 1-5 people**. Each team's member contribution should be highlighted. Different teams can definitely choose the same project.
- **Written Report:** You will write a **maximum 4 pages** PDF report on your findings using LaTeX, but we will also welcome the inclusion of an additional Jupyter notebook where you illustrate how you reproduced the main results. In general, please make sure to take into account the following considerations:

1. *Reproducibility*: Your classmates should be able to reproduce your results based on your report only. Describe what pre-processing is required, what hyper-parameter values you selected and why, and clearly describe the overall pipeline you used.
2. *Code*: Try to make your code as readable as possible. Instructions on how to run your code should also be provided.
3. *Clarity*: Explain clearly what you are trying to do. Explain why you do it. The reader should understand your work **without** having to read the original paper.
4. *Critical thinking*: If you do not reproduce similar results as in the original paper, be critical of your approach, and (perhaps) of the one in the paper.

- A LaTeX template that could be used for your report is given here.
- If the authors of the paper provided code, you may use it, but note that you will, then, be expected to go beyond just running their code, and instead apply it to new examples or settings.
- In some cases, you will find that you cannot reproduce some results due to computational constraints e.g. used datasets are too big. In those cases, feel free to try reproducing the results using a simpler or smaller dataset.
- We expect you to be creative and critical of your work and the work of others. You should try to form your own opinion on the reproduced work.
- We are definitely aware that not all tasks have the same difficulty. We will take the difficulty of the question into account when grading, so none of the choices will have any disadvantage for you.

Project 1. Reconciling modern machine learning practice and the bias-variance trade-off.

This work, by *Mikhail Belkin, Daniel Hsu, Siyuan Ma, Soumik Mandal*, showed that the generalization performance as a function of the number of parameters of a machine learning method can sometimes look very different from its classical presentation. In particular, the authors describe what they call the "double descent" phenomenon, demonstrating that over-parameterisation does not necessarily hurt generalization in machine learning. Many plots can be easily reproduced. We suggest Fig. 2 and Fig. 4, but many more are equally interesting.

Project 2. Understanding deep learning requires rethinking generalization.

This paper, by *Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, Oriol Vinyals*, was critical in showing that the standard wisdom regarding generalization was completely wrong at the time: by training neural networks with random labels, the authors debunked many classical beliefs about generalization, and sparked a new fruitful generation of works which critically studied generalization in deep learning. All figures from this paper are interesting and should/could be reproduced. However, you may find that it could be better to use simpler dataset, e.g., CIFAR or Fashion MNIST, rather than Imagenet!

Project 3. Towards learning convolutions from scratch.

This paper, by *Behnam Neyshabur*, argues that radically changing the type of optimizer can imbue a simple fully connected neural network with the right inductive biases to classify complex images. Table 2 and Fig. 3 and Fig. 4 are of great interest

Project 4. Overcoming catastrophic forgetting in neural networks.

This paper, by *James Kirkpatrick et al.*, studies the setting of *continual learning*, where an agent learns multiple tasks sequentially, i.e., accumulating knowledge for a new task/experience without having access to data from past tasks. In this setting, the phenomenon of *catastrophic forgetting* arises; the performance of the agent on past experiences significantly diminishes upon learning new tasks. The paper proposes the use of a regularization term to combat this problem. We suggest steering from the reinforcement learning experience and, instead, focus on supervised learning. Reproduce the results of Figure 2 (subfigures A + B) on the following datasets: `PermutedMNIST` (as in the paper) as well as `RotatedMNIST` (ten tasks of rotating `MNIST`, where task i is produced by a fixed rotation of $10(i-1)$ degrees). Compare the results of the proposed regularization with naive L2 regularization and no regularization.

Project 5. Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics.

This work, by *Alex Kendall, Yarin Gal, Roberto Cipolla*, proposes an adaptive scheme based on homoscedastic uncertainty to weigh the task losses in the setting of multi-task learning. Reproduce the results of Fig. 2 and Table 1. The paper experiments on the computationally expensive `CityScapes` dataset. For this reason, we propose to instead reproduce the results on two datasets: `MultiMNIST` dataset (2 tasks) and `CelebA` dataset[1] (pick 3 out of the 40 tasks).

[1]As a model, use the ResNet variant and the experimental setting of this paper, as outlined in Fig. 7 and Appendix C.2.