Practical Course - Analysis of new phenomena in machine/deep learning

Introduction Meeting

Technical University of Munich

Department of Informatics

Outline

Machine learning and deep learning research

- Empirical studies, providing benchmark and demonstrating pitfalls.

- Rigorously explain why ML / DL works by analysing theoretical models or algorithms.

Machine learning and deep learning research

- Empirical studies, providing benchmark and demonstrating pitfalls.
- Rigorously explain why ML / DL works by analysing theoretical models or algorithms.

Focus:

- Insights for new algorithmic development (example: boosting, methods for regularisation).
- Brings concepts from mathematics to ML (example: Random graphs, Geometry).

Machine learning and deep learning research

This Practical:

- Understand recent advances
- Reproduce existing results
- Extend research (empirically)

Course Setup

Basics

Setup

- 1 Paper per person
- Groups of two for discussion (but graded individually)

Basics

Setup

- 1 Paper per person
- Groups of two for discussion (but graded individually)

Main Parts

- First half of the semester: Reproduce the empirical results
- Second half of the semester: Extending the experiments (or theory)
- End of the semester (exact time will be announced): Final Presentation

Weekly Schedule

In groups of 6 students (split by supervisor: Mae, Han, Maha):

- 1h Weekly presentation. 5 min. per student + 5 min. Q&A
- 1h Office hour

Agree with your supervisor on a time.

Evaluation Format - Reproducibility Report

- Deadline roughly mid semester (2nd or 3rd week of June)
- One Jupyter Notebook
  - Readme
  - One code / plot block for each reproduced part
  - Max 300 words markdown each

Evaluation Format - Final Report

- Deadline end of the semester (dates will be announced later)
- Deadline for final report on extensions roughly two weeks after presentations so you can incorporate feedback

Evaluation Format - Final Report

- Deadline end of the semester (dates will be announced later)
- Deadline for final report on extensions roughly two weeks after presentations so you can incorporate feedback
- Report (latex template will be given) - one page for each extension
- Jupyter Notebook for the additional experiments / plots
  - Readme
  - One code / plot block for each reproduced part
  - Max 300 word markdown each

Grading

- Report on reproducibility (40%)
- Report on extensions (20%)
- Final presentation (40%)

Code

- Push code to practical Git (access will be given later)
- Repository is also used for submitting reports
- Everyone will have access to an LRZ server instance for the course (Instructions on Moodle)

Reproducibility - common questions

- What does Reproducibility mean?— If we have plots reproduce plots

Reproducibility - common questions

- What does Reproducibility mean?— If we have plots reproduce plots

- Do you need to reproduce all plots?— No. The idea is to reproduce the **main** results.

Reproducibility - common questions

- What does Reproducibility mean?— If we have plots reproduce plots
- Do you need to reproduce all plots?— No. The idea is to reproduce the **main** results.
- What if there are no plots?— If we only have theoretical bounds under specific planted settings add numerical simulations that illustrate the bounds

Reproducibility - common questions

- What does Reproducibility mean?— If we have plots reproduce plots
- Do you need to reproduce all plots?— No. The idea is to reproduce the **main** results.
- What if there are no plots?— If we only have theoretical bounds under specific planted settings add numerical simulations that illustrate the bounds
- What if there is code online?— You are allowed to use the code. In most cases the authors provide the code repository. **Discuss about the code quality and challenges you faced in the report.** Reproducibility is to check the **main idea** by varying experiments (e.g. is the trend still the same with different regularization? is the non-linearity important?)

Reproducibility - common questions

- What does Reproducibility mean?— If we have plots reproduce plots
- Do you need to reproduce all plots?— No. The idea is to reproduce the **main** results.
- What if there are no plots?— If we only have theoretical bounds under specific planted settings add numerical simulations that illustrate the bounds
- What if there is code online?— You are allowed to use the code. In most cases the authors provide the code repository. **Discuss about the code quality and challenges you faced in the report.** Reproducibility is to check the **main idea** by varying experiments (e.g. is the trend still the same with different regularization? is the non-linearity important?)

For all the above check with your supervisor if your plan is sufficient.

Possible Topics

Formal Verification for Transformer (RASP)

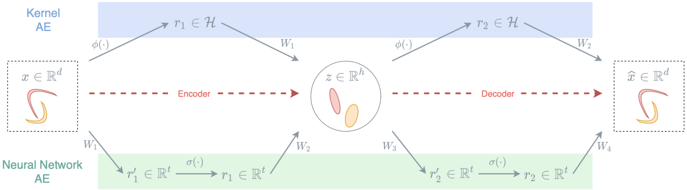
RASP (Restricted Access Sequence Processing Language) is a computational model for Transformers. It allows computational problems (e.g., acceptance of languages like k-Dyck) to be encoded in a program, which can then be compiled into an equivalent Transformer model. Ensuring the **safety** (i.e., does not crash) and the **correctness** (i.e., performs the intended functionality) of the RASP program is crucial.

- **Abstract Specification:** What specifications must each operation satisfy to ensure correctness?
- **SMT solving:** How can SMT solvers verify the program returns the desired result, considering all possible branches?
- **Conversion to Transformer*:** Can RASP be extended to return attention matrices and weights of feed-forward layer beyond just heat maps?

Scalable Kernel Representation Learning

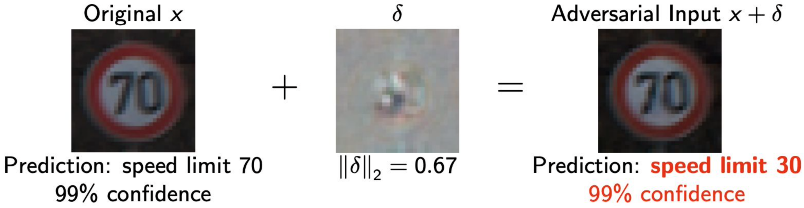
- Kernels provide a principled way to perform non-linear learning
- relying on functional analytic foundations
- Provide interpretability

We explore how one could build kernel-based foundation models by scaling kernel methods, thus enabling them utilize self-supervised approaches to learn meaningful representations.



Adversarial ML / Robustness

- Performance of NNs significantly affected if data is slightly perturbed.
- Why? How can we build robust ML models / guarantee robustness?


Original x
$+$
δ
$=$
Adversarial Input $x + \delta$
Prediction: speed limit 70
99% confidence
$\|\delta\|_2 = 0.67$
Prediction: speed limit 30
99% confidence

Paper Assignment (Also on Moodle)

Paper Assignment

- List of papers is published in Moodle
- Give your preferences by **Friday, 25.04.2025**
- Mention the following:
  Study program: Bachelor or Master
  Semester:
  Preferences: submit at least 5 preferences (ex. 5, 10, 13)

Questions

- What if my group member drops out?— No problem. Since the grading is individual you can continue without any changes.

Online Form

Please also fill in the following formular: https://forms.gle/LmhxJhtbVCWJH8LU8.

Figure: Scan Me!