Prof. Dr. Jasmin Blanchette
Dr. Martin Desharnais-Schäfer
Dr. Michael Kirsten
Elisabeth Lempa

Ludwig-Maximilians-Universität München
Institut für Informatik
Discussion on 14.01.2026
Homework due on 21.01.2026 at 16:00

Possible solution for Exercise Sheet 12 in

Scientific and Technical English for Computer Scientists

The exercise sheets consist of in-class exercises and homework. The in-class exercises take place during the second half of the lecture time slots. The homework, which is optional and ungraded, can be submitted via the "Homework" section in Moodle. The homework is subject to peer review.

Unless indicated otherwise, generative artificial intelligence assistants such as ChatGPT may be used, as long as you acknowledge how you use them as specified by the Institute's policy on plagiarism.[1] However, you may not use such tools to generate peer reviews for you. In addition, we strongly recommend that you do not use them to generate entire solutions, since that would defeat the purpose of the exercises.

In-class exercise 12-1 ***Slide Competition*** Among the material available on Moodle, you will find the three slide decks for these presentations:

1. Richard Yuanzhe Pang, Weizhe Yuan, Kyunghyun Cho, He He, Sainbayar Sukhbaatar, and Jason Weston, "Iterative Reasoning Preference Optimization," *NeurIPS 2024*, 2024.[2]

2. Konstantin Fischer, Ivana Trummová, Phillip Gajland, Yasemin Acar, Sascha Fahl, and M. Angela Sasse, "The Challenges of Bringing Cryptography from Research Papers to Products," *USENIX Security 2024*, 2024.[3]

3. Patrizio Angelini, Therese Biedl, Markus Chimani, Sabine Cornelsen, Giordano Da Lozzo, Seok-Hee Hong, Giuseppe Liotta, Maurizio Patrignani, and Sergey Pupyrev, and Ignaz Rutter, and Alexander Wolff, "The Price of Upwardness," *GD 2024*, 2024.[4]

a) Have a look at them and briefly note what you consider their respective strengths and weaknesses.

[1] `https://www.medien.ifi.lmu.de/lehre/Plagiate-IfI.pdf`
[2] `https://neurips.cc/media/neurips-2024/Slides/96659_nZhhoGi.pdf`
[3] `https://www.usenix.org/system/files/usenixsecurity24_slides-fischer.pdf`
[4] `https://ac.tuwien.ac.at/gd2024/slides/Session10-Patrizio-Angelini.pdf`

b) Discuss the slide decks and their respective strengths and weaknesses in groups of two or three, and rank the decks from strongest to weakest according to your collected arguments. Which arguments are the most important to you? Can you reach a consensus in your group?

POSSIBLE SOLUTION:

First Slide Deck

Strengths:

- Pleasing color scheme
- Well-readable fonts and spacing of text
- Highlighting and slide reveal clarify mapping from figures to text
- Highlighting and slide reveal identify clear flow of the presentation
- Slides not too cluttered and not too many slides

Weaknesses:

- No examples
- Many abbreviations
- Measurements without units

Second Slide Deck

Strengths:

- Well-readable fonts
- Not too many slides
- Slide reveal supports flow of presentation
- No unclear abbreviations

Weaknesses:

- Some very cluttered slides, especially figure and summary
- Inconsistent text spacing (e.g., items on slide 22)

Third Slide Deck

Strengths:

- Well-designed animations for clear flow of presentation
- Well-used colors for illustrations of graph properties
- Illustrations not too complex
- Sparing use of abbreviations, most of them introduced
- Not too many slides

Weaknesses:

- Some quite cluttered slides, but mitigated by slide reveals
- Some long text fragments buried within illustrations

Most Important Arguments

Clear flow, simple examples, few abbreviations.

Ranking

Third, then second, then first.

In-class exercise 12-2 ***Poster Presentation*** Among the material available on Moodle, you will find this scientific poster:

> Mayee F. Chen, Nicholas Roberts, Kush Bhatia, Jue Wang, Ce Zhang, Frederic Sala, and Christopher Ré, "Skill-it! A Data-Driven Skills Framework for Understanding and Training Language Models," *NeurIPS 2023*, 2023.[5]

a) Have a look at the poster and identify its domain and its main takeaway.

POSSIBLE SOLUTION:

Domain: Machine learning and language models

Main takeaway: Language models are more efficient when stratified for distinguished learning skills.

b) Pretend that you are one of the poster's authors, and develop a short script for a one-minute presentation of your poster to a professor who is an expert in the poster's domain.

[5] `https://neurips.cc/media/PosterPDFs/NeurIPS 2023/72098.png`

POSSIBLE SOLUTION:

Large language models are trained on massive datasets, but this is expensive and often inefficient. Inspired by human learning, we hypothesize that models acquire skills in a structured order. To capture this, we define a skill as a set of training examples tied to a specific capability, and build a directed skill graph where edges represent prerequisite relationships. If training skill A improves performance on skill B, we add a directed edge from A to B.

Our main contribution is Skill-It, an online sampling algorithm that uses this graph to prioritize training examples based on how much each skill influences downstream tasks. Skill-It dynamically adjusts its sampling strategy, creating an adaptive curriculum.

On synthetic and real-world benchmarks such as "LEGO synthetic skills" and RedPajama, Skill-It outperforms uniform sampling, improving generalization and reducing computational cost. Our work offers a new framework for understanding language model learning and designing more efficient, interpretable training strategies.

c) Develop a slightly longer script for a two-minute presentation of the poster to a doctoral researcher who has just started working in the poster's domain.

POSSIBLE SOLUTION:

Language models such as GPT are trained on massive amounts of text, but not all data is equally useful. Language models are only as good as the data on which they are trained. Uniform training is not efficient—it is like teaching a student random topics without a plan.

Hence, we draw inspiration from how humans build knowledge through interconnected skills, ordered by difficulty. For example, in math, we learn addition and subtraction before multiplication, and only then move on to algebra.

So we ask: "Can we train models more efficiently by organizing the data around the skills they need to learn?"

We define a skill as a group of examples that teaches the model a specific capability—like answering questions, translating, or summarizing. But skills are not isolated—many depend on others. Like in math, some skills are prerequisites of others. To capture this, we introduce the skill graph, where nodes are skills and edges represent dependencies.

With the skill graph, we develop Skill-It, an algorithm that helps train models more efficiently. Instead of randomly sampling data, Skill-It chooses

which skills to focus on by evaluating

i) how important each skill is to the final objective;

ii) how well the model is already performing on that skill.

Skill-It actively samples data that helps the model improve where it needs it most. This creates a structured, adaptive learning process.

In our experiments, Skill-It substantially improved accuracy and reduced validation loss on synthetic tasks. On real-world data, it achieved better performance with less training data.

So the big idea is this: If we understand how models acquire skills over time, we can guide their learning, making training faster, cheaper, and more effective.

d) Read both of your scripts to a neighbor, and have them read their scripts. Discuss your presentations, comparing them with respect to style and focus.

POSSIBLE SOLUTION:

For the first script, I focused on the technical structures, the algorithm, and the data sets. For the second script, I focused on illustrating the main concepts and the underlying analogy, and emphasized the main takeaways.