Third workshop on Metascience for Machine Learning

Hayley Hung, Marco Loog, Jan van Gemert
Dec 5 2025

ellis | UNIT DELFT

What is Metascience for Machine Learning?

Holding a magnifying glass up to the ways of doing ML research

Metascience for machine learning focuses on the science in the field of machine learning. It's about topics that are typically not found in Machine Learning text books.

It's about asking how and why research in machine learning is done in certain ways, the methodology, the processes, the mindset, goals, aspirations, inspirations.

At its heart, metascience for ML seeks to critically evaluate, understand, and enhance the way we generate, assess, and communicate scientific knowledge in machine learning—guiding the field toward more rigorous, reflective, and meaningful progress.

Goals of Metascience for ML

1. Scoping:
What is metascience for machine learning
formalised via written (joint) perspectives

2. Strategic Development:

What activities could be helpful to address the problem?

What are the low hanging fruit,

where are opportunities for additional seed funding/catalysers?

3. Who and What:

who is available to contribute what in the developed strategy?

Making the "what" **desirable and manageable**

ELLIS Delft Funded Activities by the end of 2025

- Workshops

  - **June 20** : Pitches, Develop separate subgroups / task forces (potentially new collaborations) that would work on specific assignments.

  - **Oct 9** : metascience of benchmarking/ continue task force discussions

  - **Dec 5**: metascience in other disciplines/ task force updates and plans. (10:00-14:30)

    - Daniel Lakens TUEindhoven, Abigail Delgado, UUtrecht

- Web page and social media

- 2 Podcasts:

  - Is this a comfortable and appropriate medium to reach the masses?
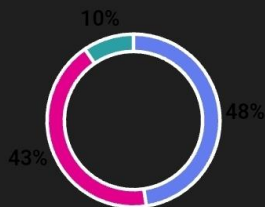
Summary of Activities so far

Third Workshop Today:

- ~22 signups

4. Select your role

More details

- Faculty staff (Assistant, Associate, Full Professor) — 10
- Temporary staff (PhD or Postdoc) — 9
- Other — 2

10%
48%
43%

Podcast

- **First Pilot Recording**
  - **June 9th VMB, floor 6, 12:00-13:00 with live audience.**
  - Topic: What is Metascience for Machine Learning?
- **Second recording**
  - **Friday Oct 31 13:30-15:00 with live audience**
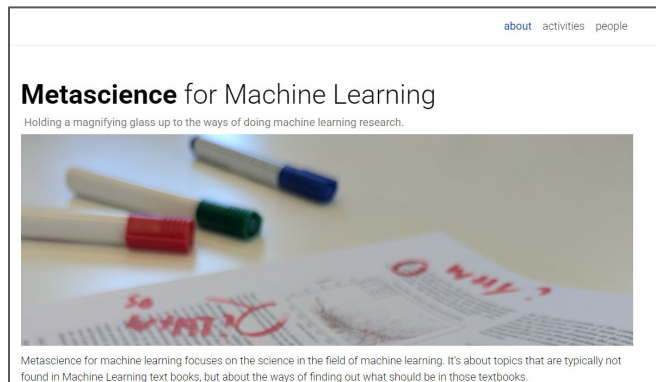 - **Topic: neurips is hampering scientific progress in ML**
- **Future recordings bi monthly**
  - Aiming for… Lunchtimes on a Wednesday
  - Record 5 propositions
  - Improving podcasting recording setup.

Summary of Activities so far

Web page

https://metascienceforml.github.io/

Metascience for Machine Learning

Holding a magnifying glass up to the ways of doing machine learning research.

Metascience for machine learning focuses on the science in the field of machine learning. It's about topics that are typically not found in Machine Learning text books, but about the ways of finding out what should be in those textbooks.

People page needs more people

| June 2025 | We organized our first workshop on metascience for machine learning on June 20, 2025! ✨ |
| March 2025 | We got a starting fund from Ellis unit Delft 😊 |

Working groups?

We will find out more today!

New research theme in MSC DSAIT

New Mailing list

Schedule

- **10:00: Walk in with coffee/tea**

- **10:10: Welcome (Hayley)**

- **10:20: Presentation by Abigail Nives Delgado (University of Utrecht),**
- **about "Philosophy of Science as a Tool for Scientific Practice"+ discussion**

  Chair: Marco

- **11:15: Presentation by Daniel Lakens (TU Eindhoven),**
- **about "How to Improve Our Science: Lessons from the Replication Crisis in Psychology"+ discussion**

- **12:10: Lunch**

- **13:00: post lunch reflection (talks, activities, nextsteps?)**

- **13:20 Brief update subgroup activity [based on prior reporting]**

  Chair: Hayley

- **13:30: Plenary discussion on new topics/activities/directions**

- **13:35: Break out discussions**

- **14:15: Plenary summary**

  Chair: Jan

- **14:35: Closing**

- **14:45: End**

House Rules

- **Be mindful of sub-disciplinary culture:**
  - what you consider basic knowledge may be entirely unknown to your conversation partner.
- **Be prepared to play devil's advocate**
- **Avoid drowning in details:**
  - Identify key problems at the abstract level
- **Be aware of our stakeholders:**
  - ELLIS, ML community, ….?
- **Be mindful that strategy is implemented by concrete action:**
  - For each abstracted problem, have
    - short term easy goals (2025)
    - longer term more ambitious goals 2026-...?

Brief explanation/inventory of existing working groups and proposals of new ones; incentives people may need to put some serious effort into MS4ML

Group AA - Julian, Copilot, et al.

Explore and exploit middle ground between theory and empiricism

- Noted lack of framework to describe diversity in research problems and approaches
- Need for more theory of experimental research

Group 25B - Rickard et al.

Measuring the problem in MS4ML [scientific rigor, "influencers", reproducibility,...]

- 3 meetings to discuss various projects that could turn into bachelor or msc projects. Targetting q4 and msc next year

Group Next - Tiffany et al.

Guidelines [for good practice?]

- First steps to gather existing guidelines for review
- Next : discussion on what these guidelines should set out to do
- Continue today in next session?

Group AA : Bridging Theory and Empiricism in ML

Short-Term (within 1 month?)

- Clarify scope and structure of the two proposed taxonomies through internal discussion.
- Draft initial outlines for both taxonomies, including candidate categories and dimensions.
- Identify key literature and existing frameworks that could inform taxonomy development.
- Define criteria for "landmark tasks" and begin collecting candidate examples.
- List assumptions commonly made in ML papers, especially in applied domains.

Open Questions for the Group

- What criteria should we use to define a "landmark task" in ML?
- How can we distinguish between robust empirical regularities and overfitted heuristics?
- What would a good experimental test of a theoretical ML claim look like?
- How can we incentivize the publication of negative results or failed replications?
- Should we aim to formalize empirical theories, or is it enough to document and test them?
- How can we ensure that taxonomies remain useful and not overly rigid or prescriptive?

Brief explanation/inventory of existing working groups and proposals of new ones; incentives people may need to put some serious effort into MS4ML

- How do we make sure we actually follow through?

- One approach : come up with agreed-upon rule / expectation for how group members hold themselves and each other responsible for their commitments and results

Breakout

- Appoint a note-taker [somebody else could chair]
  - Be ready to share your top project, working agreement, and timeline

- Define 1 to 3 projects with clear scope and desired outcome
- State at least one key working agreement for sustaining commitment
- Outline the first concrete steps for your top project
- Set a initial-deliverable date

      [note the hyphen; aim for one month from now]

            …and set a target completion date