

COMP64803: Introduction to Responsible & Reproducible AI

Mingfei Sun

CDT on Decision-Making for Complex Systems
Department of Computer Science
The University of Manchester



In the last lecture

We've learned how to use:

- ▶ Remote development
- ▶ Computational Shared Facility (CSF)
- ▶ Local workstations (4090 RTX)

In this lecture

- ▶ Paper Reproducibility Challenge
- ▶ Project Workflow
- ▶ Expected Outcomes
- ▶ Available Resources

Paper Reproducibility Challenge (1/3)

To produce reproducibility reports on very recent papers published at

- ▶ Eleven top ML conferences:

NeurIPS 2024, ICML 2024, ICLR 2024, ACL 2024, EMNLP 2024, CVPR 2024, ECCV 2024, AAAI 2024, IJCAI-ECAI 2024, ACM FAccT 2024, SIGIR 2024,

- ▶ Or top ML journals in 2024/2025

JMLR, TACL and TMLR

Why reproducibility is important?

Essentially, think of your role as an inspector verifying the validity of the experimental results and conclusions of the paper.

- ▶ To encourage the publishing and sharing of scientific results that are reliable and reproducible
- ▶ To verify the empirical results and claims in the paper by reproducing the computational experiments
- ▶ Either via a new implementation or using code/data or other information provided by the authors

Your role will also extend to helping the authors improve the quality of their work and paper.

Paper Reproducibility Challenge (3/3)

How our reproducibility challenge work?

You do not need to reproduce all experiments in your selected paper:

- ▶ For example the authors may experiment with a new method that requires more GPUs than you have access to
- ▶ You could elect to reproduce only the baseline results
- ▶ You can check if baseline methods are properly implemented
- ▶ You can check whether hyper-parameter search is done with the same degree of attention

Paper Reproducibility Challenge (3/3)

How our reproducibility challenge work?

You do not need to reproduce all experiments in your selected paper:

- ▶ For example the authors may experiment with a new method that requires more GPUs than you have access to
- ▶ You could elect to reproduce only the baseline results
- ▶ You can check if baseline methods are properly implemented
- ▶ You can check whether hyper-parameter search is done with the same degree of attention

If available, the authors' code can and should be used

- ▶ The methods described can also be implemented/re-implemented according to the description in the paper
- ▶ You are supposed to detect anomalies in the code, or shed light on aspects of the implementation that affect results.

Project Workflow (1/2)

1. Select a paper and avoid duplicate work (due by 24th March EoD)
 - ▶ Encourage to avoid duplicate reproducibility efforts: search for the papers with open-sourced codes
 - ▶ Or papers that have been reproduced by other reproducibility challenges
2. Submit form with the selected papers and codes you find (due by 24th March EoD)
 - ▶ A form will be live on our Teams channel
 - ▶ Title, authors, conference, link to the paper, link to the code
 - ▶ Expectations from your supervisors

Project Workflow (2/2)

3. Post reproducibility project on our GitHub wiki pages (due by 7th April EoD)
 - ▶ Outline the reproducibility plan, including the experiments you plan to reproduce and milestone dates
 - ▶ Report your initial critics on the paper and the code
4. Submit a report (due by 30th June EoD, one month after the examination period)
 - ▶ This report should have an abstract (300-400 words) detailing the key findings of your report
 - ▶ Your report should be in ICLR Latex template, within 5-8 pages.
 - ▶ Your report should include assessment comments from your supervisors (as a separate section titled "Supervisor's Assessment")

Timeline

Mar 2025	SEM2W05	24						1	2
	SEM2W06	25	3	4	5	6	7	8	9
	SEM2W07	26	10	11	12	13	14	15	16
	SEM2W08	27	17	18	19	20	21	22	23
	SEM2W09	28	24	25	26	27	28	29	30
	SEM2W10	29	31	Submit paper selection					

Apr 2025	SEM2W10	29		1	2	3	4	5	6
	EVAC 01	30	7	8	9	10	11	12	13
	EVAC 02	31	14	15	16	17	18	19	20
	EVAC 03	32	21	22	23	24	25	26	27
	SEM2W11	33	28	29	30				

May 2025	SEM2W11	33				1	2	3	4
	SEM2W12	34	5	6	7	8	9	10	11
	Ex SEM2W13	35	12	13	14	15	16	17	18
	Ex SEM2W14	36	19	20	21	22	23	24	25
	Ex SEM2W15	37	26	27	28	29	30	31	

Jun 2025	Ex SEM2W16	37						1	
	SVAC 01	38	2	3	4	5	6	7	8
	SVAC 02	39	9	10	11	12	13	14	15
	SVAC 03	40	16	17	18	19	20	21	22
	SVAC 04	41	23	24	25	26	27	28	29
	SVAC 05	42	30	submit your report					

Expected Outcomes

- ▶ You should produce a Reproducibility report, describing the target questions, experimental methodology, implementation details, analysis and discussion of findings, conclusions on reproducibility of the paper.
- ▶ The result of the reproducibility study should NOT be a simple Pass / Fail outcome. The goal should be to identify which parts of the contribution can be reproduced, and at what cost in terms of resources (computation, time, people, development effort, communication with the authors).
- ▶ You should expect to engage in dialogues with the paper authors.
- ▶ The top reports are encouraged to be submitted to <https://reproml.org/>

Available Resources

- ▶ 4090 RTX workstations (will be reserved for this reproducibility challenge for the next 3 months)
- ▶ Computational Shared Facility (CSF)
- ▶ Your supervisors (discuss with your supervisors)

END LECTURE