

Reproducibility budgets for ML preprints

Example Author A

Example Author B

Example Author C

2026-05-12

Demonstration paper in the rrxiv reference corpus. The canonical machine-readable version lives at rrxiv.com/papers/rrxiv:2605.00003.

Abstract

We argue that preprint platforms should require authors to declare a reproducibility budget at submission time — an explicit specification of the compute, data, and human-time required to replicate the paper’s central claims. We evaluate adoption costs over 12 months and report mixed results: smaller labs benefit, large labs resist.

1 Introduction

We argue that preprint platforms should require authors to declare a reproducibility budget at submission time — an explicit specification of the compute, data, and human-time required to replicate the paper’s central claims. We evaluate adoption costs over 12 months and report mixed results: smaller labs benefit, large labs resist.

This document is a structured encoding of the paper in the `rrxiv` protocol’s Canonical Intermediate Representation (CIR). It engages with the topics `stat.ML` and `cs.LG`. The encoding registers 3 formal claims (1 contradicted, 1 partial, 1 replicated). Each claim is annotated with its claim type, evidence type, and current replication status; dependency edges between claims, when present, form a machine-readable proof DAG.

2 Methodology

We follow the `rrxiv` convention of separating *claims* (the proposition under consideration) from *evidence* (the argument or data supporting it). Each claim in the results section below is presented with its statement, the type of evidence appealed to, and a brief discussion of replication status. Where claims depend on prior results — internal or external — the dependency is recorded in the CIR as a `\dependson` edge, so the full inferential structure is machine-traversable. Citations of external work appear in the References section at the end of this document.

3 Results: registered claims

Claim 1

Claim 1 (Claim 1). Reproducibility-budget declarations correlate with 2.3x higher replication success rates ($p < 0.05$).

Replication status: replicated.

This claim is an empirical observation supported by data. As of the encoding date, it has been independently replicated.

Claim 2

Claim 2 (Claim 2). Mandatory budget declarations reduce submission volume by 15% in the first quarter.

Replication status: contradicted.

This claim is an empirical observation supported by data. As of the encoding date, it has been contradicted by subsequent work.

Claim 3

Claim 3 (Claim 3). Budgets self-correct via community feedback over a 6-month window.

Replication status: partial.

This claim is an empirical observation supported by data.

4 Discussion

The claim graph above is the primary product of this paper. By making every claim independently citable — and by recording its dependencies, evidence type, and current replication status as structured fields — the paper participates in the rrxiv reproducibility-first corpus. Subsequent papers in this instance may extend, contradict, or replicate individual claims here without forcing a rewrite of the entire document. See the canonical version online for the live discourse layer.

5 References

- Computational reproducibility at scale
- Reproducibility in machine learning