

A Year of "AI" Shenanigans

SSI Fellows Update

Jesper Dramsch (2022 Cohort)



Tackling the looming reproducibility crisis in science
that uses machine learning in a world consumed by AI
hype.

The 2022 Plan

- 🤝 Participate in Community
- 🚧 Create a workshop
- 🎥 Do YouTube
- ✒ Write articles
- 🔊 Give Talks

Slight Diversion: A tutorial

- Reproducibility in ML for Science
- 6 Jupyter Notebooks
- EuroScipy
- "Hiding the Brokkoli"

by Jesper Dramsch





- Increase citations, ease review & foster collaboration

Bad Conscience: A Workshop

- Recruit fellow fellows
- Write proposal in days
- Get accepted way late
- Make an awesome workshop anyways
- 2 hours with 70 people

by Jesper Dramsch



realworld-ml.xyz

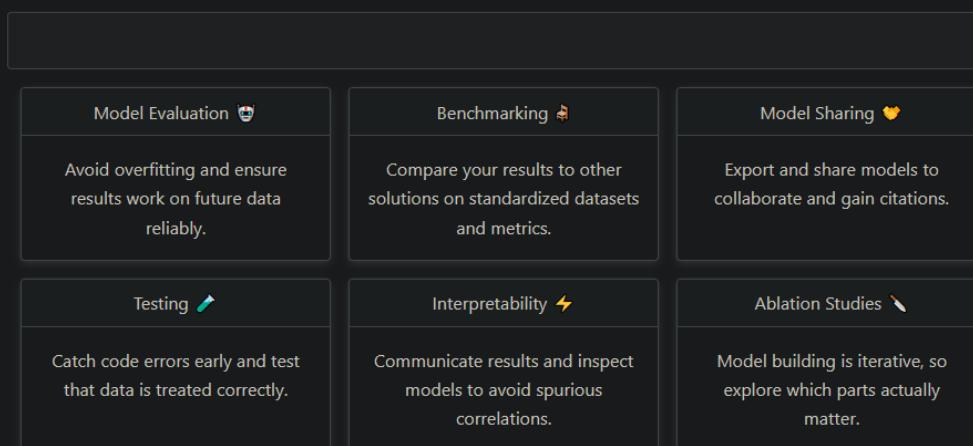
A collection of "easy wins" to make machine learning in research reproducible.

This book focuses on basics that work. Getting you 90% of the way to top-tier reproducibility.

Every scientific conference has seen a massive uptick in applications that use some type of machine learning. Whether it's a linear regression using scikit-learn, a transformer from Hugging Face, or a custom convolutional neural network in Jax, the breadth of applications is as vast as the quality of contributions.

This tutorial aims to provide easy ways to increase the quality of scientific contributions that use machine learning methods. The reproducible aspect will make it easy for fellow researchers to use and iterate on a publication, increasing citations of published work. The use of appropriate validation techniques and increase in code quality accelerates the review process during publication and avoids possible rejection due to deficiencies in the methodology. Making models, code and possibly data available increases the visibility of work and enables easier collaboration on future work.

This work to make machine learning applications reproducible has an outsized impact compared to the limited additional work that is required using existing Python libraries.



This book is organized into these major sections:

- **Motivation** to expand on how the following sections aide in increasing citations, easing review, and fostering collaboration.
- **Front Matter** that goes into the installation and data.
- **How To** with notebooks and additional resources on the sections to improve research artifacts.
- **Talks & Workshop** that showcase presentations around this material.

Overall, this tutorial is aimed at applied scientists that want to explore machine learning solutions for their problems.

Take it all the way

- Make Tutorial accessible and appealing
- Jupyter Book
- Extra Content
- Visually Appealing
- Search-enginey

by Jesper Dramsch

ML.Recipes

Other stuff

- Reproducible ML at work
- Taught MOOC to 7,000 people
- Skillshare courses with ethical and reproducible AI
(6,500 total students)
- 5 SSI and several of mine Blog posts
- 20 Youtube Videos with 1.4M impressions
- Guest on multiple Podcasts
- Multiple talks and guest lectures

Budding Community

- Machine Learning
- But make it inclusive
- 🏳️🌈-friendly
- Neurodivergent
- Do my best on everything else

The Latent Space

latent.club

Have fun with your
Fellowship

Embrace the Chaos!