An Interactive Demo of Multiplicity in ML

Prakhar Ganesh[1,2]

[1]McGill University, [2]Mila - Quebec AI Institute
prakhar.ganesh@mila.quebec

Goal of the Demo

There is a growing debate about the implications of *multiplicity*—conflicting behavior among a set of "good" models (Marx, Calmon, and Ustun 2020)—for algorithmic decision-making. This demo introduces the phenomenon of multiplicity in machine learning and highlights associated concerns, ranging from conflicting predictions (Black, Raghavan, and Barocas 2022; Gomez et al. 2024) to risks of homogenization (Creel and Hellman 2022; Jain et al. 2024).

Prerequisite Knowledge and Target Audience

The demo is designed to be accessible to a multidisciplinary audience. No technical prerequisite knowledge is required, though a basic familiarity with how machine learning models are trained is assumed.

Expected Engagement Time

The demo is self-paced, but on average, we expect it to take 8-10 minutes to complete.

Setup Instructions

The demo is built using the *'streamlit'* library in *python*. It was originally designed on Huggingface Spaces and is currently hosted at the following link,

https://huggingface.co/spaces/prakharg24/multiverse

Running Locally. The demo can also be set up locally. All relevant files and instructions are provided. Refer to the *'README.txt'* file for further instructions.

Brief Description of the Material

The demo is divided into six interactive pages, with guiding text in between. We provide a brief overview of all six pages.

Page 1: Information Loss

On the first page, the user creates their own data by selecting features to collect for a loan application approval model. While the user selects from ideal features (for example, 'financial discipline'), the demo then shows what can be realistically collected (for example, 'credit score'). This gap between the real world and the collected data helps emphasize the information loss central to any data-driven learning.

Page 2: Rashomon Effect

The user is shown a toy dataset and several models that all give the same accuracy (known as the Rashomon set). The user is then given the option to choose one model, and based on their choice, the demo highlights an individual who received a negative prediction but would have received a positive prediction if a different choice had been made (i.e., multiplicity). This visualizes multiplicity and emphasizes the impact of developer choices on individuals.

Page 3: Implicit Developer Decisions

Continuing with the same toy dataset, the user is then asked to make decisions during model development, instead of directly choosing the final model. However, the decisions lead to the same models as before, showing how decisions during development impact the final trained model.

Page 4: List of Developer Decisions

This page allows the user to understand how many different decisions are made during model development. At the top, the decisions are divided into five categories (Data Collection, Data Processing, Model Selection, Model Training, Model Evaluation). When the user clicks on a category, five more sub-categories are presented. And when the user clicks on a sub-category, a set of developer decisions are presented.

Page 5: Intention-Convention-Arbitrariness (ICA)

Given the scale of developer decisions, it is not possible to ensure each decision is intentional. On this page, we visualize the intention-convention-arbitrariness (ICA) triangle by Ganesh, Taik, and Farnadi (2025). The user can change the 'amount' of intentionality, conventionality, or arbitrariness in a decision, leading them to various parts of the triangle, which highlights unique examples in those regions.

Page 6: Multiverse

The final page of the demo is a complete machine learning pipeline, with several decisions being left to the user. The demo visualizes how these decisions together can explode the number of possible models, allowing the users to navigate through this 'multiverse' by making different decisions.

References

Black, E.; Raghavan, M.; and Barocas, S. 2022. Model multiplicity: Opportunities, concerns, and solutions. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 850–863.

Creel, K.; and Hellman, D. 2022. The algorithmic leviathan: Arbitrariness, fairness, and opportunity in algorithmic decision-making systems. *Canadian Journal of Philosophy*, 52(1): 26–43.

Ganesh, P.; Taik, A.; and Farnadi, G. 2025. Systemizing Multiplicity: The Curious Case of Arbitrariness in Machine Learning. In *Proceedings of the 2025 AAAI/ACM Conference on AI, Ethics, and Society*.

Gomez, J. F.; Machado, C.; Paes, L. M.; and Calmon, F. 2024. Algorithmic Arbitrariness in Content Moderation. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, 2234–2253.

Jain, S.; Suriyakumar, V.; Creel, K.; and Wilson, A. 2024. Algorithmic Pluralism: A Structural Approach To Equal Opportunity. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 197–206.

Marx, C.; Calmon, F.; and Ustun, B. 2020. Predictive multiplicity in classification. In *International Conference on Machine Learning*, 6765–6774. PMLR.