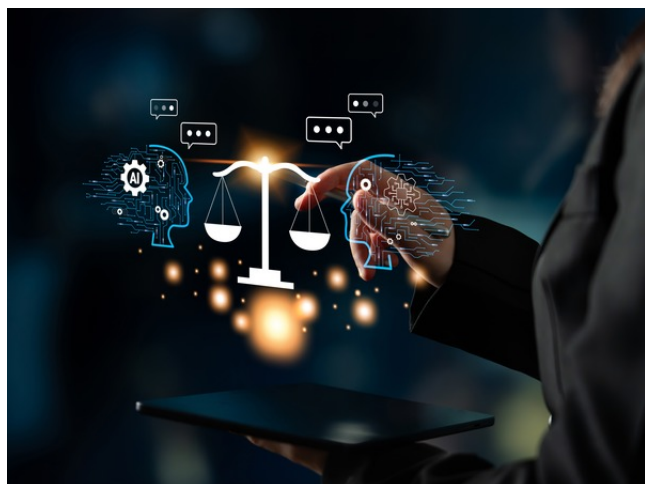




Implementations of Generalization Gap Error Predictors

TECHNOLOGY NUMBER: 2024-249



OVERVIEW

Provides validated codebase for evaluating generalization gap error predictors under simplicity bias

- Streamlines testing and comparison of GEPs with standardized, bias-aware evaluation tools
- Machine learning research, model benchmarking, algorithm development, educational purposes

BACKGROUND

In machine learning, understanding how well models generalize from training to unseen data is crucial. The "generalization gap"—the performance difference between these two phases—has spurred the development of metrics and predictors to estimate it early. However, most prior evaluations of generalization error predictors (GEPs) lacked standardized benchmarks, and few addressed the influence of "simplicity bias," a tendency for models or predictors to favor simple solutions. This oversight can skew GEP effectiveness and hinder reproducible scientific progress. Without unified, high-quality reference implementations and a robust test-suite, researchers face challenges in fairly assessing new approaches and making meaningful advancements toward more reliable, generalizable machine learning models.

INNOVATION

Technology ID

2024-249

Category

Software

MOSS - Michigan Open Source
Support

Inventor

Danai Koutra

Puja Trivedi

Jayaraman Thiagarajan

Further information

Ashwathi Iyer

ashwathi@umich.edu

[View online](#)



This work introduces a comprehensive, open-source codebase containing high-quality implementations of state-of-the-art generalization gap error predictors (GEPs) along with a dedicated test-suite for assessing their performance under simplicity bias. The provided suite enables researchers to consistently benchmark new predictors and systematically explore the effects of simplicity bias, delivering clear comparisons across different methods. By consolidating validated GEP routines and standardized evaluation practices, this innovation accelerates research reproducibility and method improvement within the machine learning community. Real world applications include advancing model evaluation science, informing the design of more robust learning algorithms, supporting algorithm selection in industry, and serving as an educational resource for machine learning curricula.

ADDITIONAL INFORMATION

PROJECT LINKS:

- [Github](#)

DEPARTMENT/LAB:

- [Danai Koutra, Electrical Engineering and Computer Science \(EECS\)](#)

LICENSE: