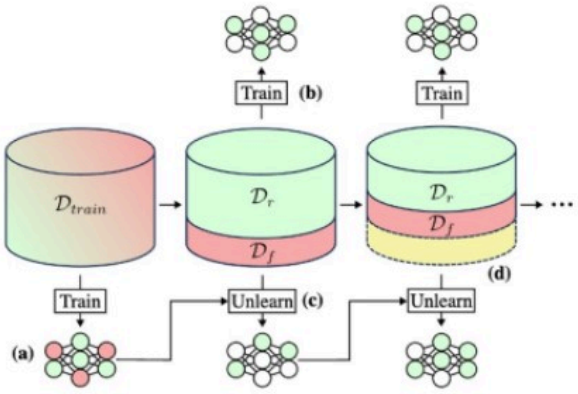
SEI Bulletin

Trouble reading this email? View in browser.



3 Recommendations for Machine Unlearning Evaluation Challenges

August 28, 2024—The integration of machine learning models into more products and services has raised concerns about controlling the data a model is trained on. But retraining a model from scratch to remove specific data points is often impractical. Research into machine unlearning is developing new methods to remove data points from a model efficiently and effectively. A new SEI Blog post discusses machine unlearning methods and recommends ways to evaluate them more robustly.

"Current machine unlearning methods are limited in their measured effectiveness and scalability," write the authors of *3 Recommendations for Machine Unlearning Evaluation Challenges*. "We're working on developing new machine unlearning evaluations that more accurately mirror a production setting and subject models to more realistic privacy attacks."

Read the blog post »

SEI News

SEI Launches Course on Developing a National or Government Computer Security Incident Response Team
The new course is intended to help nations develop robust cybersecurity capabilities.

Ozkaya Explores Generative AI in Software Architecture During International Conference Keynote
The IEEE International Conference on Software Architecture is the premier event for the field's researchers and practitioners.

See more news »

Latest Blogs

3 Recommendations for Machine Unlearning Evaluation Challenges
Machine unlearning aims to remove data points efficiently and effectively from a model without the need for extensive retraining. This post describes our recommendations for robust evaluations of machine unlearning methods.

Building Quality Software: 4 Engineering-Centric Techniques
Why is it easier to verify the function of a software program rather than its qualities? Alejandro Gomez outlines four engineering-centric techniques to creating quality software.

See more blogs »

Latest Podcasts

3 API Security Risks (and How to Protect Against Them)
McKinley Sconiers-Hasan discusses three API risks and how to address them through the lens of zero trust.

Evaluating Large Language Models for Cybersecurity Tasks: Challenges and Best Practices
Jeff Gennari and Sam Perl discuss applications for LLMs in cybersecurity, potential challenges, and recommendations for evaluating LLMs.

See more podcasts »

Latest Videos

Generative AI and Software Engineering Education
SEI experts in software engineering discuss how generative AI is influencing software engineering education.

Secure Systems Don't Happen by Accident
Timothy A. Chick discusses how security is an integral aspect of the entire software lifecycle.

Latest Publications

Using Quality Attribute Scenarios for ML Model Test Case Generation
This conference paper presents an approach based on quality attribute (QA) scenarios to elicit and define system- and model-relevant test cases for machine learning models.

Lessons Learned in Coordinated Disclosure for Artificial Intelligence and Machine Learning Systems
This paper describes lessons learned from coordinating AI and ML vulnerabilities at the SEI's CERT/CC and observations of public discussions of AI vulnerability coordination cases.

See more publications »

Upcoming Events

DevSecOps Days Washington D.C. 2024, September 18
This free, in-person event in Arlington, Virginia, will teach how to integrate security into DevOps practices and transform DevSecOps journeys.

International Conference on Conceptual Modeling (ER 2024), October 28-31
The SEI will host the main international forum for discussing the state of the art, emerging issues, and future challenges in research and practice on conceptual modeling.

See more events »

Upcoming Appearances

Billington CyberSecurity Summit, September 3-6
Hear the SEI's Greg Touhill, Mark Sherman, and Nathan VanHoudnos at the leading government cybersecurity summit, and visit the SEI at booth 320.

TechNet Indo-Pacific 2024, October 22-24
Visit the SEI at booth 1411.

27th Annual Systems & Mission Engineering Conference, October 28-31
Visit the SEI booth at this event.

See more opportunities to engage with us »

Upcoming Training

Insider Risk Management: Measures of Effectiveness
October 29-31 (SEI Live Online)

Risk Program Development - Governance and Appetite Workshop
November 13-14 (SEI Arlington, Va.)

See more courses »

Employment Opportunities

Senior AI Security Researcher

Software Developer - Advanced Computing Lab

Associate Infrastructure Engineer

All current opportunities »

Carnegie Mellon University
Software Engineering Institute