PURDUE INNOVATES | Office of Technology Commercialization

A method for identifying naming mismatches in neural networks based on their architectural properties

DARA is an automated tool that uses architectural metadata to detect naming inconsistencies and potential plagiarism in Pre-Trained Models, improving discoverability, reuse, and validation.

As innovation in deep learning continues, many engineers are increasingly relying on Pre-Trained Models (PTMs) to accelerate development in deep learning systems. However, inconsistent and misleading naming practices in PTM registries, especially Hugging Face, create barriers to model discovery, reuse, and validation.

Researchers at Purdue University have developed DARA (DNN Architecture Assessment), the first automated tool to detect naming inconsistencies in PTMs using architectural metadata alone. It also includes a mixed-methods study of naming practices and user preferences across 108 Hugging Face users. Moreover, it enables automated validation, metadata generation, and plagiarism detectionâ€”functions currently missing in PTM registries.

Technology Validation:

DARA achieved a 94% accuracy in identifying model types and over 50% accuracy in detecting other architectural metadata inconsistencies.

Advantages:

-94% accuracy in model type identification

-Enables automated metadata validation and generation

-Detects naming inconsistencies and potential plagiarism

-Improves discoverability and reuse of PTMs

-Supports standardization across model registries

Technology ID

2025-DAVI-70909

Category

Artificial Intelligence & Machine Learning/AI Model Optimization & Acceleration Tools

Authors

Chingwo Cheung

James C Davis

Wenxin Jiang

Mingyu Kim

Heesoo Kim

Further information

Matt Halladay

MRHalladay@prf.org

Erinn Frank

EEFrank@prf.org

View online

Applications:

-Model zoos and PTM registries (e.g., Hugging Face)

-ML development platforms and IDEs

-AI model validation and compliance tools

-Software supply chain security

-Academic and industrial ML research

TRL: 4

Publication: https://link.springer.com/article/10.1007/s10664-025-10711-4

Intellectual Property:

Provisional-Gov. Funding, 2025-05-28, United States

Keywords: Pre-Trained Models, PTM registries, deep learning, DNN Architecture Assessment, DARA, Hugging Face, architectural metadata, model validation, plagiarism detection, software supply chain security, AI Model Reuse, Computer Technology, Deep Neural Networks (DNNs), DNN Architecture, Electrical Engineering, Empirical Software Engineering. Mixed-Methods Study, Hugging Face, Machine Learning (ML), Metadata Validation, Model Registry, Naming Consistency, Pre-trained Models, Software Reuse and Adaptation, Software Supply Chain