RESEARCH REVIEW 2022

Carnegie Mellon University
Software Engineering Institute

Maturation of Determining the Limits of AI Robustness (MDLAR)

NOVEMBER 14–16, 2022

Dr. Mike Konrad
Principal Researcher, Software Engineering Measurement and Analysis

Document Markings

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM22-0833

MDLAR: Military Application

The Department of Defense (DoD)

- Is increasingly relying on **machine learning (ML) systems** for a variety of mission and support functions
- Requires **rapid adaptation** to new situations, environments, and threats
- Needs confidence that, when deploying ML systems, they will **perform accurately**
- Needs to know **when to wait** for data to retrain/boost before their continued use

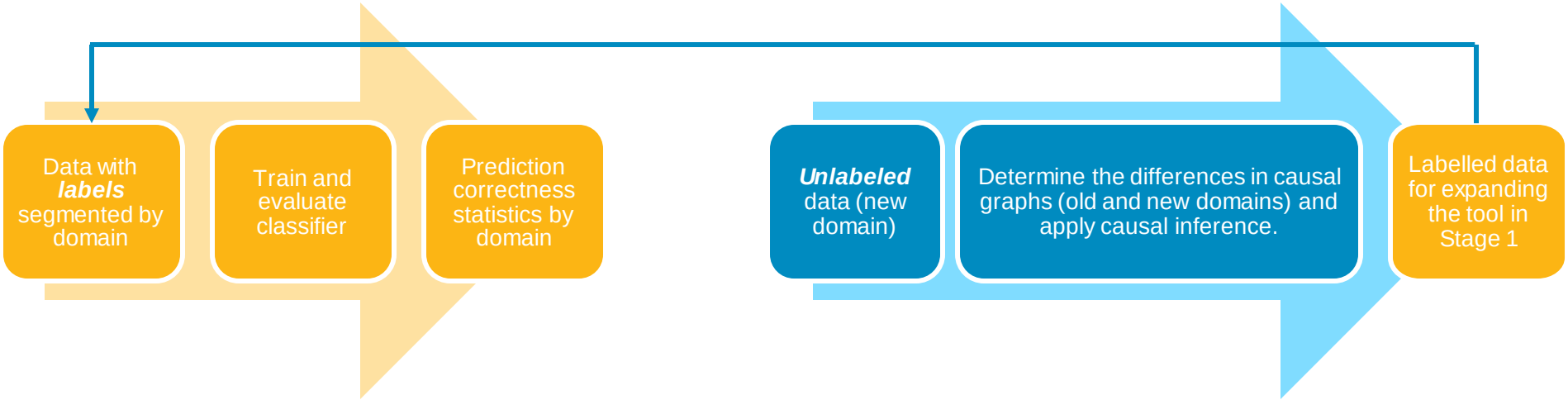
This one-year project focused on

- Developing an **approach** for determining the robustness of artificial intelligence (AI) solutions
- Building an **application** (prototype) to do just that within a narrow domain

Two Stages to Our Project

Stage 1 prototyped a secure cloud application to help an engine analyst evaluate classifiers across multiple already-labeled domains.

Stage 2 extended the solution provided in first scenario to cover unlabeled new/hypothetical domains.



Research Question. Can labels generated in this way be suitable for ML training and evaluation to cover interesting edge cases and domains?

DoD Applicability and Collaborators

Carnegie Mellon University Software Engineering Institute

Our initial **transition focus** was on organizations doing fleet maintenance that

- Collect data on **fleet performance** to maximize safe, cost-effective fleet operations
- Wonder whether AI and ML solutions are appropriate for their situation

Our transition and research **collaborators** provided demos for

- Automating the analysis of engine health fleet-wide
- Using interactive tools for evaluating classifiers on problematic edge cases

We conducted bi-weekly engagements with

- **Dr. Elias Barenboim** (Columbia University) for causal identification and estimation algorithms
- **Dr. Joe Ramsey** (CMU) for improvements needed to Tetrad

Our long-term **transition goals** are to

- Demonstrate the practicality of causal discovery and inference algorithms and toolset
- Fully automate 5-7 causal identification and estimation techniques for situational use by the DoD
- Identify transition opportunity within each DoD Service