**Motivation**

- **No guarantees** that the warnings from Static Analyzers are real bugs
- Prior work [A] has identified the Golden Features, the **most important features** from the literature
- We conduct a replication study to better understand the features

![Diagram](DIAGRAM)

**False alarms?**

- Code
- Static analyzer
- Warnings
- Classifier
- Developers

**Data leakage and duplication**

- Determine if warnings are actionable by checking the reference revision

  - Testing revision
  - Reference revision

  Collect warnings for a context

  **Data Leakage.** In computing features capturing the proportion of actionable warnings in a context, the ground-truth labels were leaked. The computation of these features require the status of warnings in the **future.**

  - Compute labels

  - Training revision
  - Testing revision
  - Reference revision

  Collect warnings

  **Duplication.** Some warnings appear in both the training and testing dataset.

  Intuition: a warning may be present before the training revision and is still present at the testing revision

  These issues are **subtle** and hard to detect, but have a **significant** impact.

**Low agreement between heuristically-determined labels and human labels**

- Compute labels

  - Training revision
  - Testing revision
  - Reference revision

  Collect warnings

  **An open warning** is present at the reference revision
  **A closed warning** is no longer present at the reference revision (but the source file is still present)

  **Heuristic:** open warnings are **false alarms**
  closed warnings are **actionable**

  However, we find that only 47% of closed warnings were actionable. Many were only closed **incidentally.**

**Roadmap**

- Data collection
- Identified the golden features

[B,C] Yang et al. (2021)
- Active Learning
- Discovered that the data is low dimensional

[D] Our work (2022)
- Investigated issues with the features
- Investigated issues with the dataset
- Features were still predictive! (AUC > 0.5)
- Motivates more work on new techniques and the need to address the lack of labelled data

[E] Yedida et al. (submitted to TSE, preprint available)
- “How to Find Actionable Static Analysis Warnings”
- “reflect more on that data”
- Open and collaboration science

**References**


[B] Yang et al. “Learning to recognize actionable static code warnings (is intrinsically easy).” EMSE 2021


[D] Kang et al. “Detecting False Alarms from Automatic Static Analysis Tools: How Far are We?” ICSE 2022

[E] Yedida et al. “How to Find Actionable Static Analysis Warnings” arxiv (currently under review)