

Can Automated Impact Analysis Technique Help Predicting Decaying Modules?



Natthawute
Sae-Lim



Shinpei
Hayashi



Motoshi
Saeki

*Tokyo Institute of
Technology*





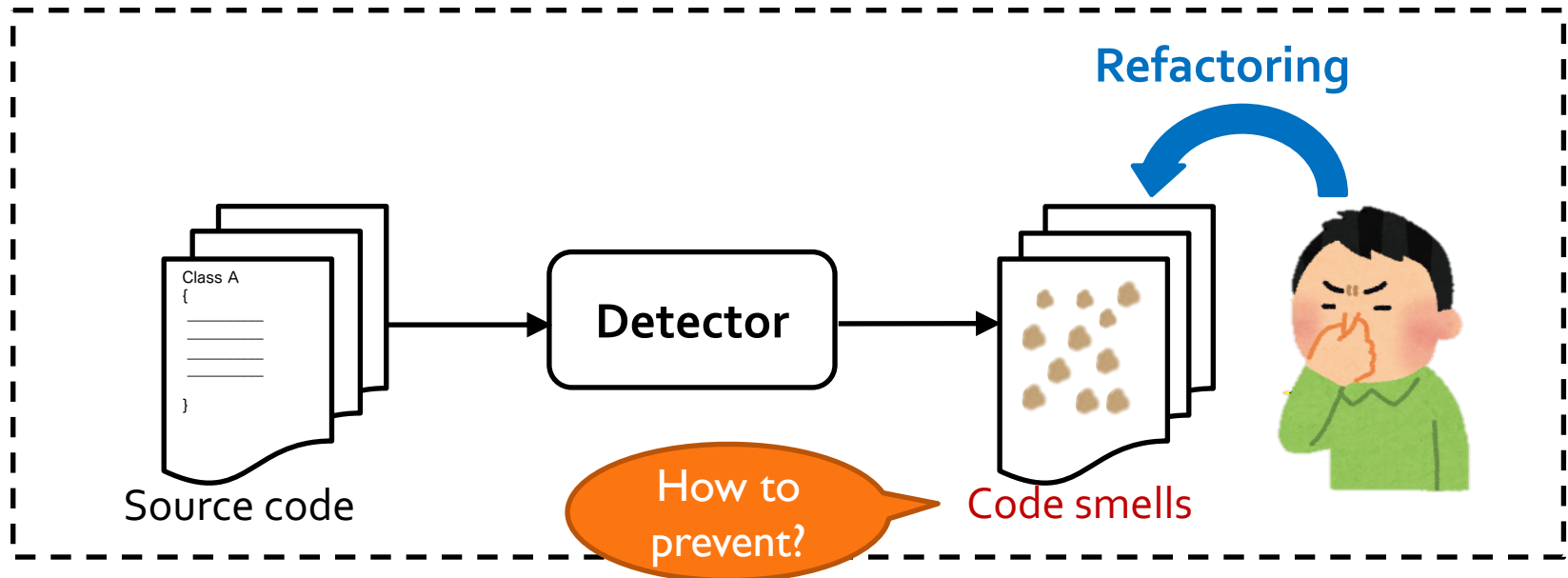
Bad Smell

Code Smell

An indicator of a design flaw/problem in the source code

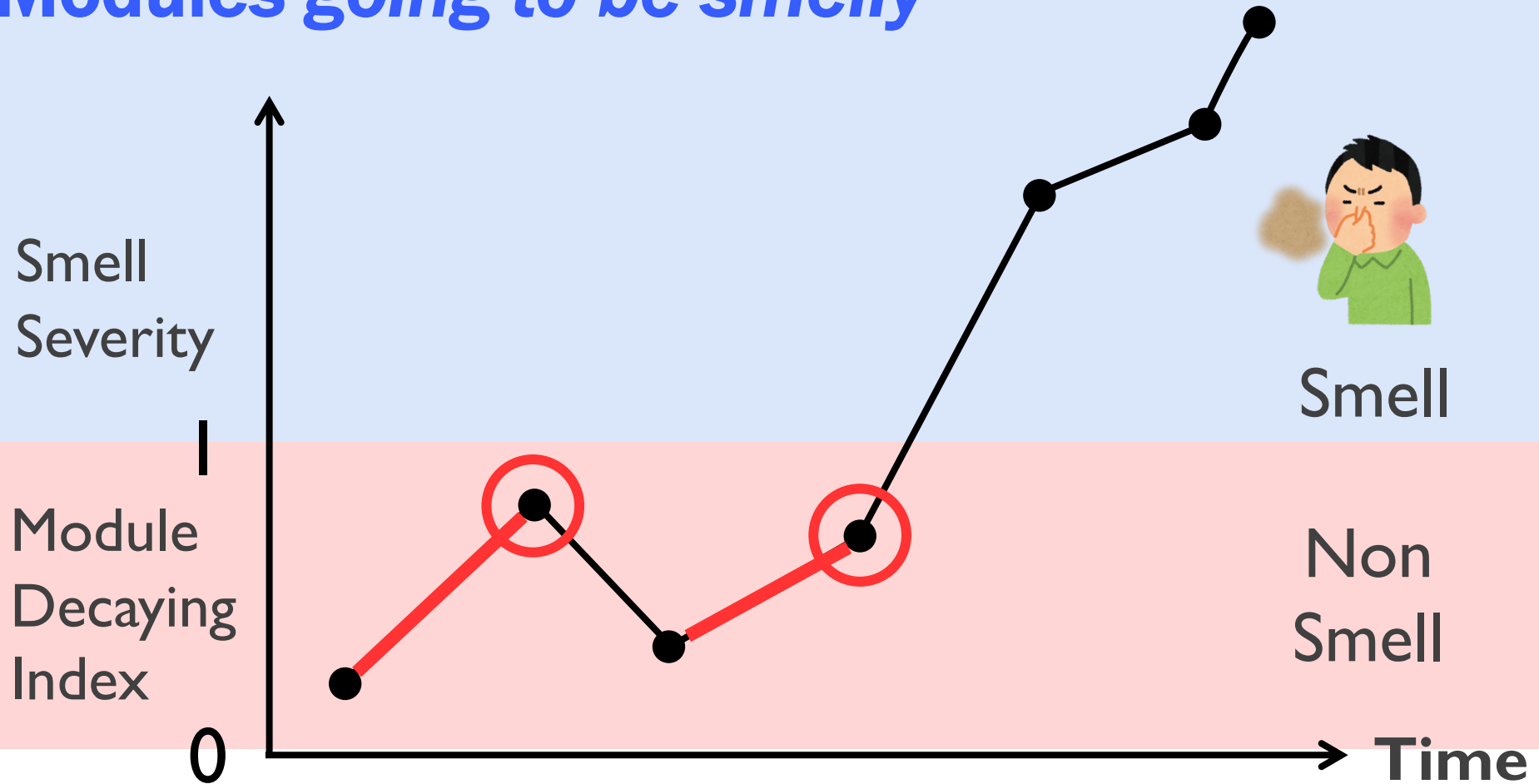
— One of the factors that cause technical debt ☹️

— Increases code component's fault-proneness ☹️



Decaying Modules:

Modules going to be smelly

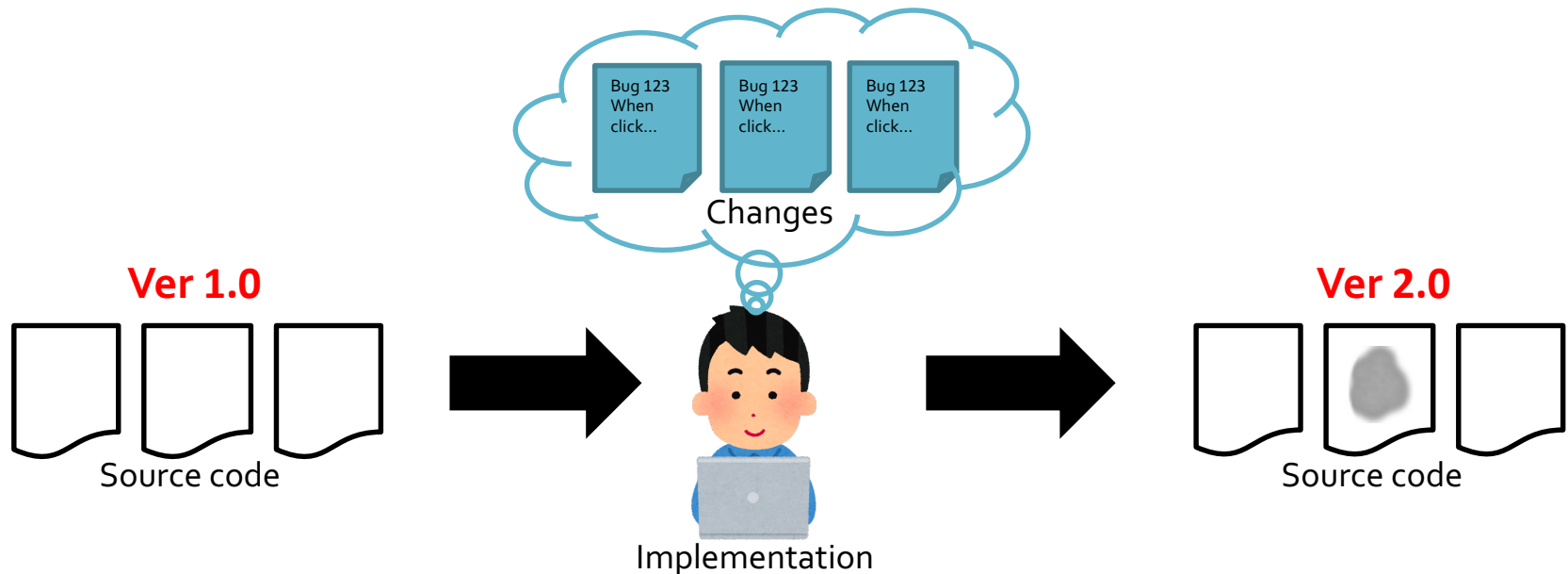


Decaying modules $\Rightarrow$ whose smell degree has increased from the previous release

- “ *approx. 19% of modules were decaying modules in each release*
- “ *decaying modules are more likely to get decayed in later releases*

Decaying Module Prediction

- Existing approach: code quality metric-based
- Use of task context
 - Using IR-based impact analysis (IA) to connect change descriptions to modules may contribute to improving the prediction performance



Research Questions

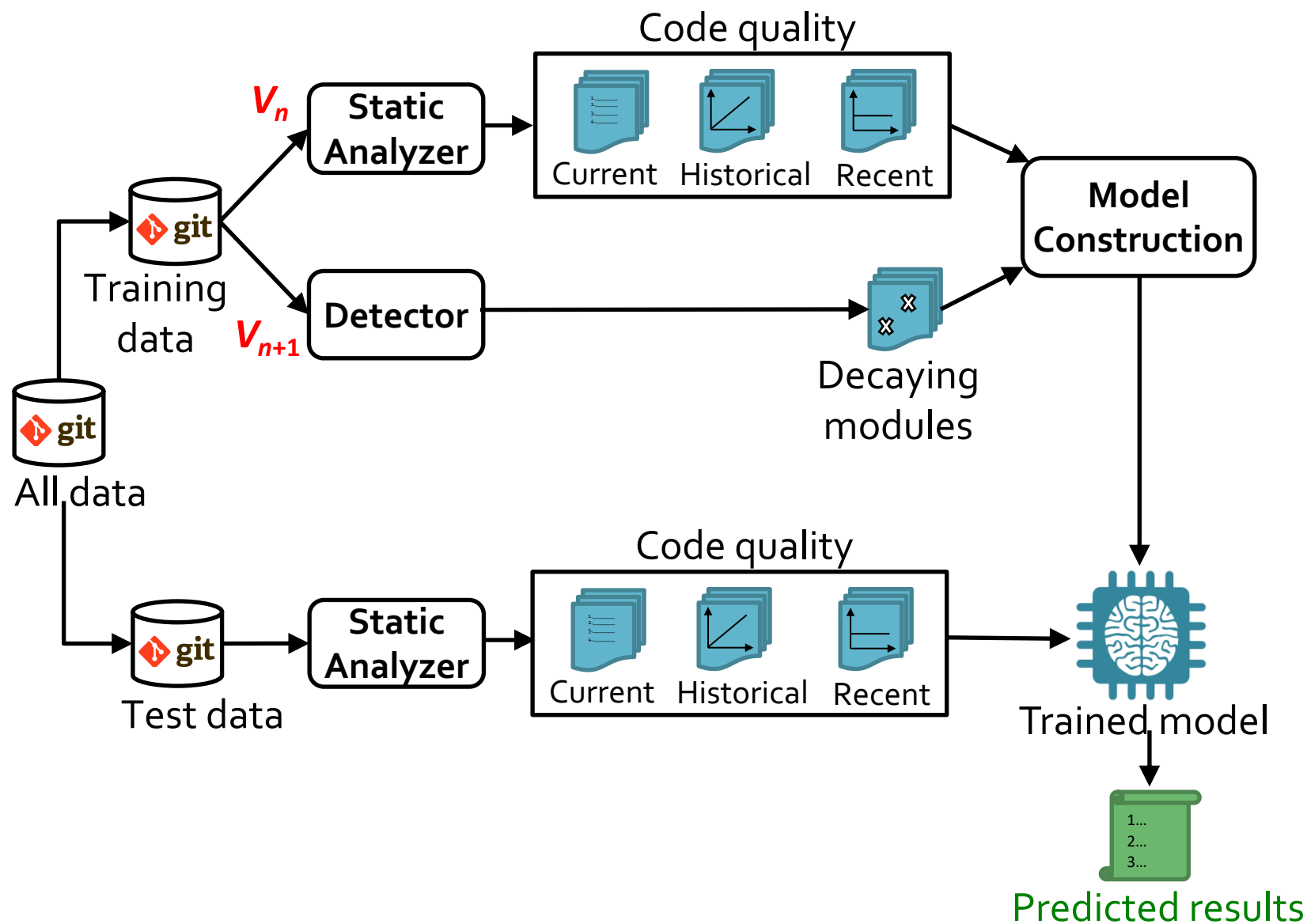
RQ1:

Can IR-based impact analysis improve prediction performance?

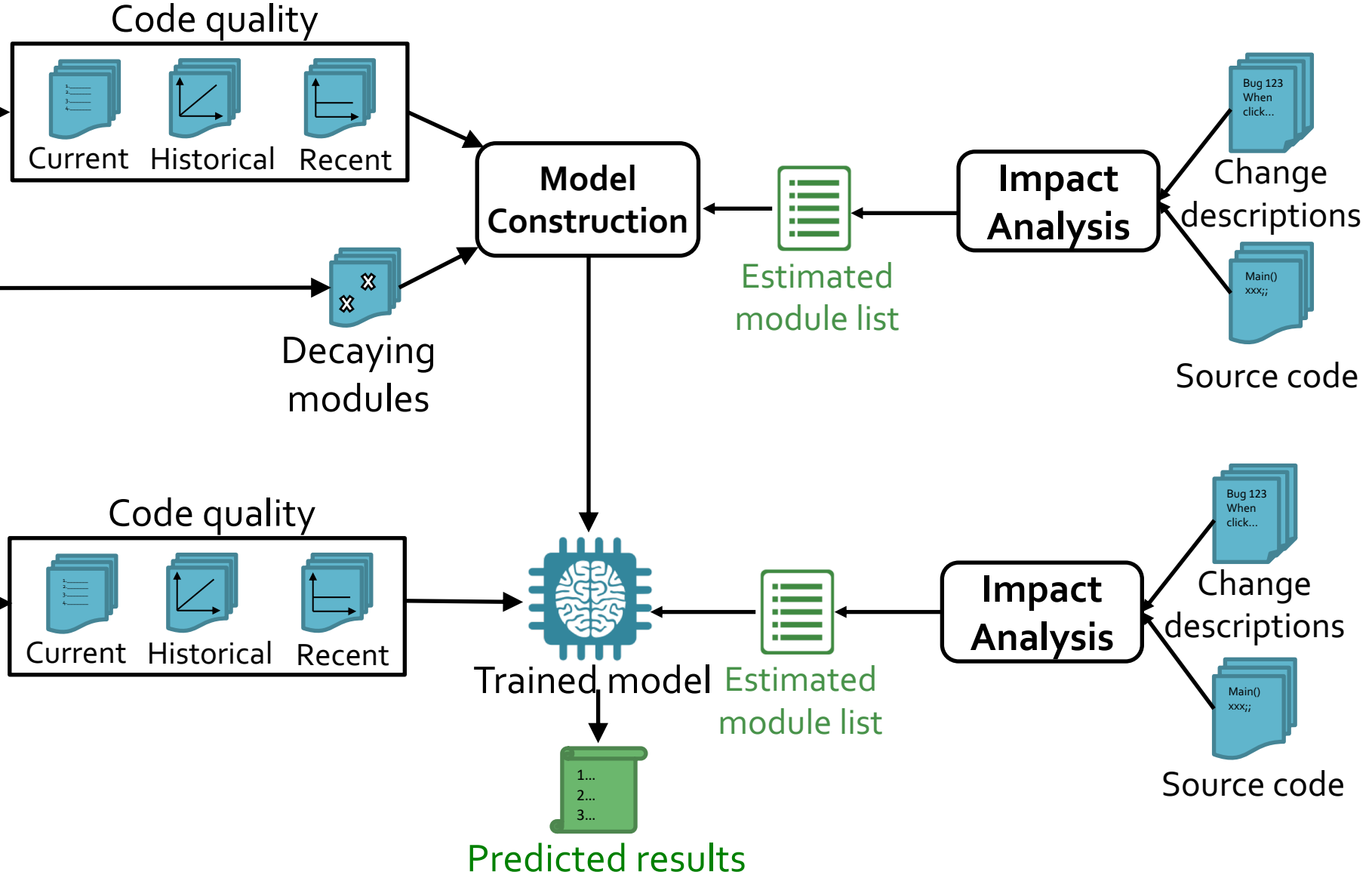
RQ2:

How can we further improve prediction performance?

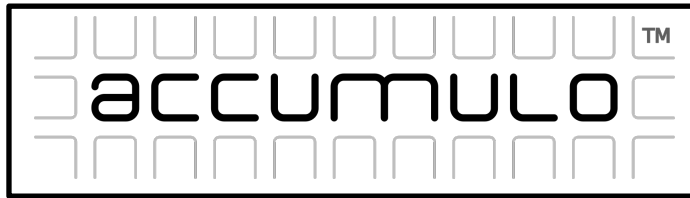
Baseline



Baseline + Impact Analysis



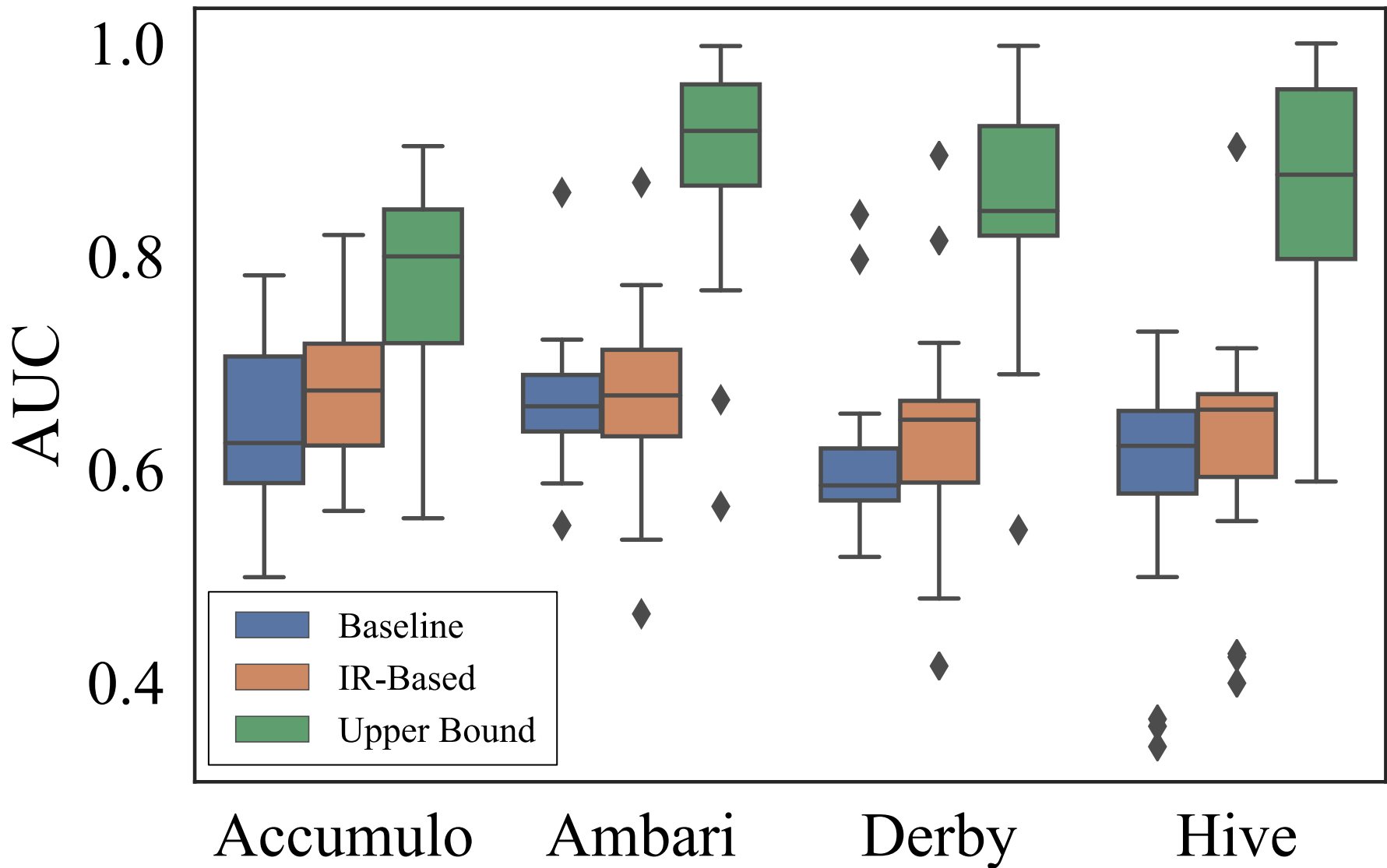
Dataset



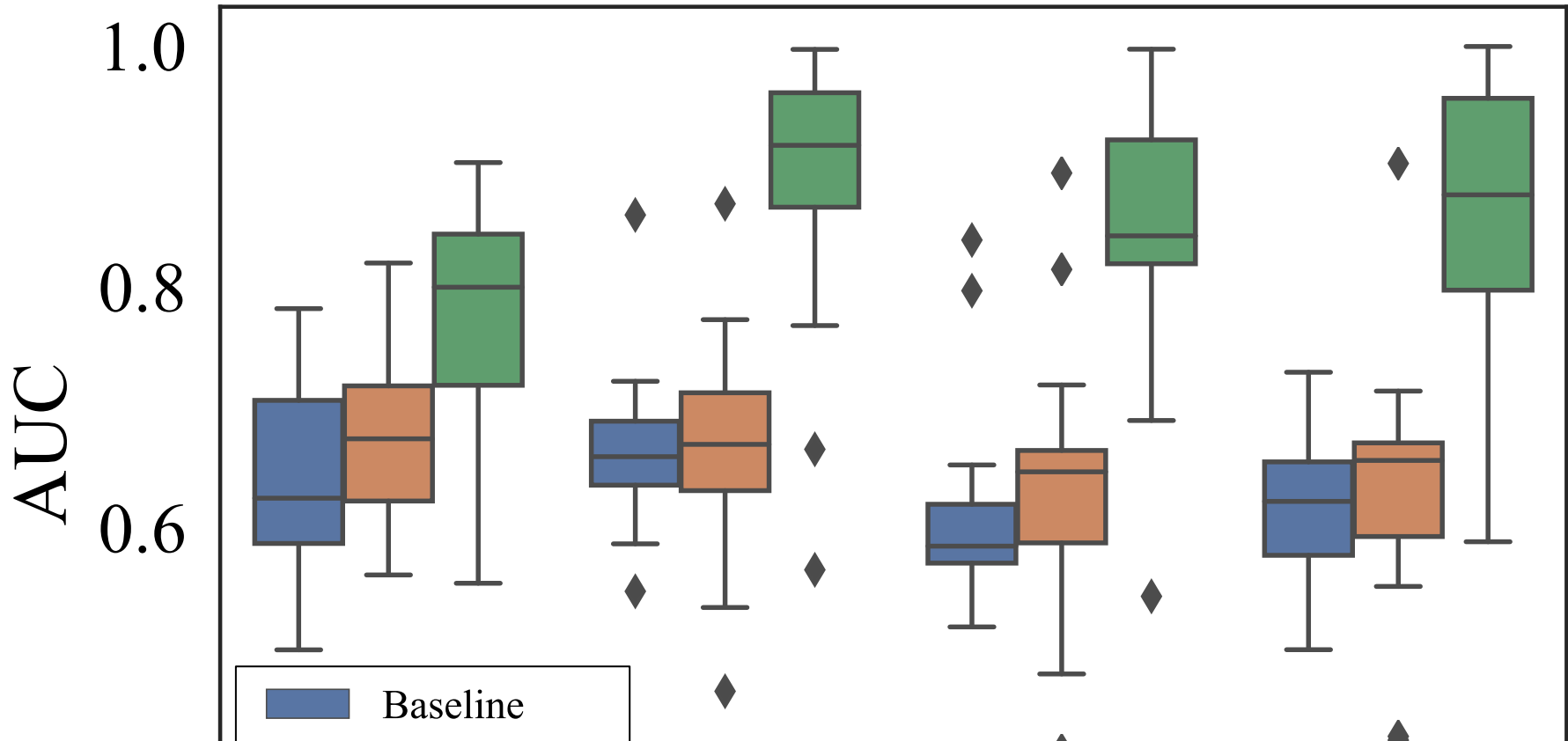
Project	Period (years)	# Releases
Accumulo	6	30
Ambari	6	33
Hive	13	28
Derby	7	33

Target smell: God Class

RQ 1 Can IR-based IA improve prediction performance?



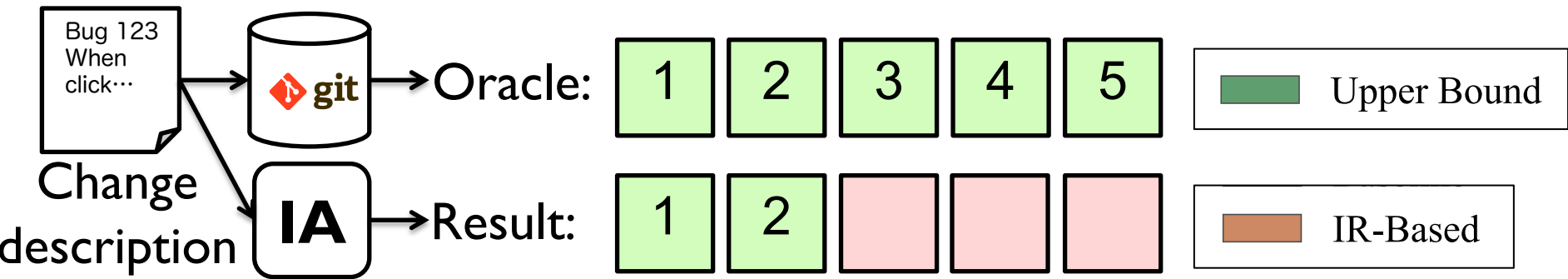
RQ 1 Can IR-based IA improve prediction performance?



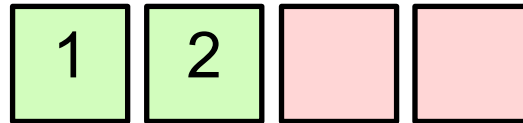
If IA is perfect, accuracy becomes very high.
Future direction: Improving IA

RQ 2 How can we further improve the performance?

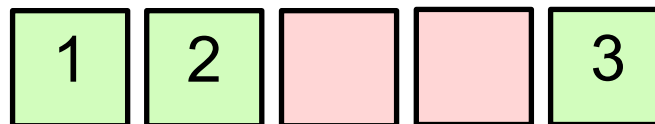
Simulating IA results of different accuracy:



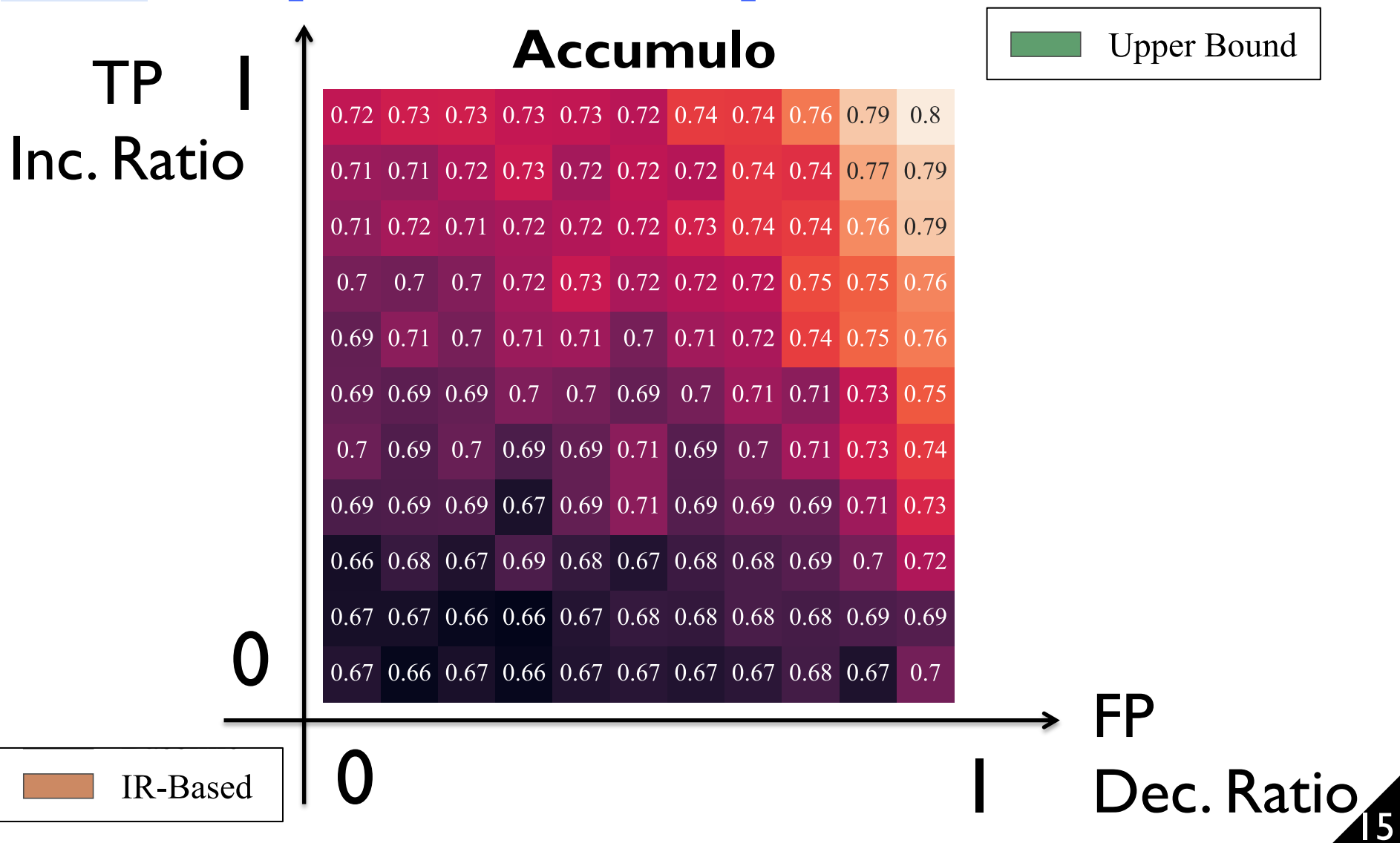
Delete ↓ false positives (FPs)



Insert ↓ true positives (TPs)



RQ 2 How can we further improve the performance?



RQ 2 How can we further improve the performance?

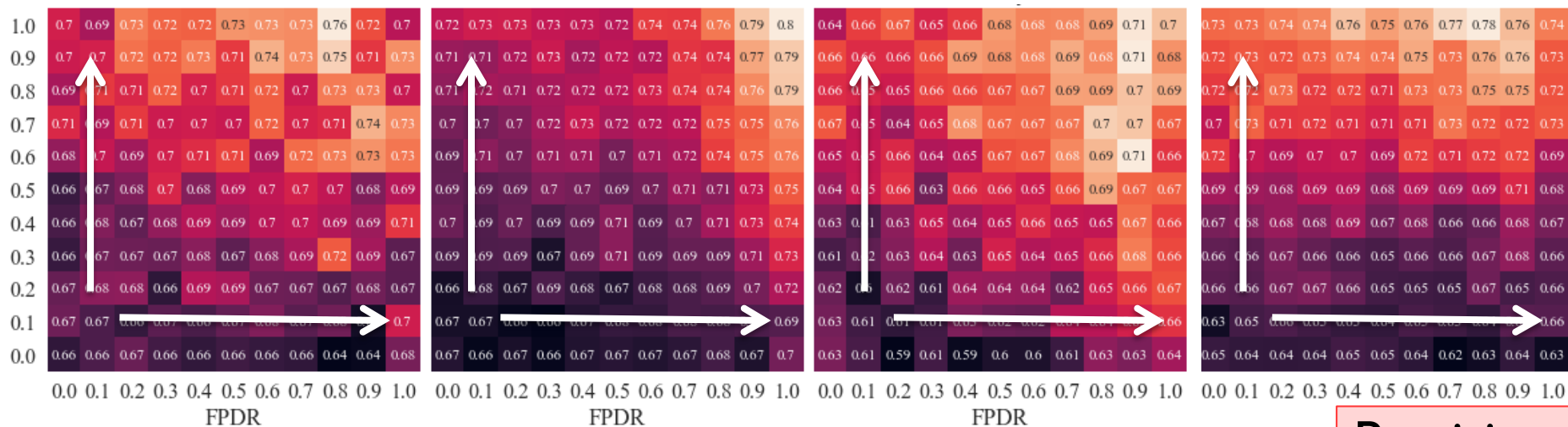
Recall

Ambari

Accumulo

Derby

Hive



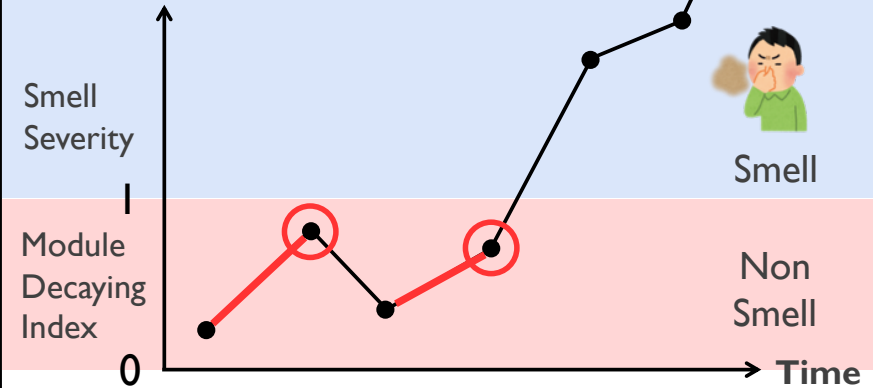
Precision

IA accuracy $\uparrow$ leads to Performance $\uparrow$

Adding TPs (recall!**) should be done first**

Can Automated Impact Analysis Technique Help Predicting Decaying Modules?

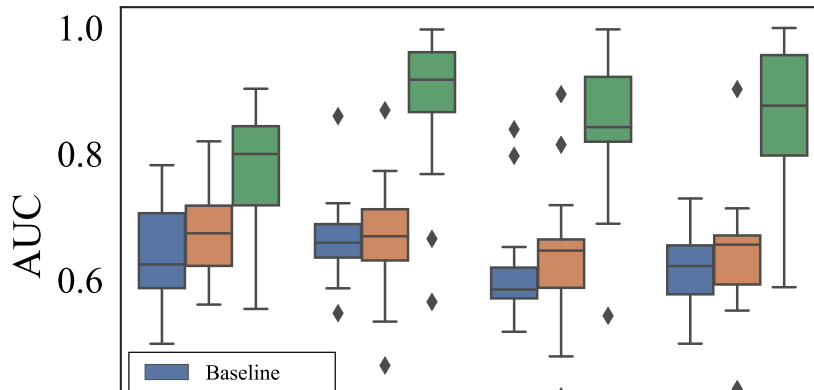
Decaying Modules: Modules going to be smelly



Decaying modules $\Rightarrow$ whose smell degree has increased from the previous release

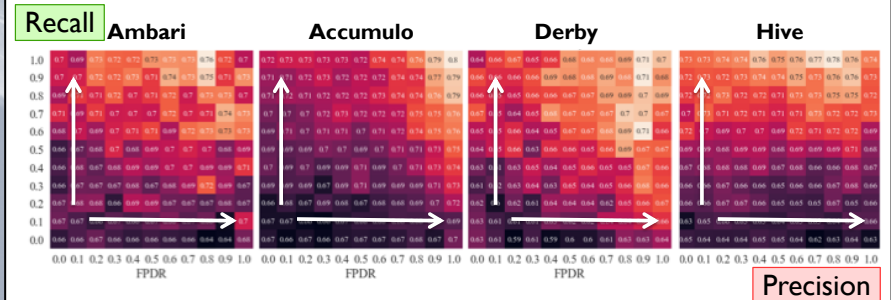
3

RQ 1 Can IR-Based IA improve prediction performance?



If IA is perfect, accuracy becomes very high
Future direction: Improving IA

RQ 2 How can we further improve the performance?



IA accuracy $\uparrow$ leads to Performance $\uparrow$

Adding TPs (recall!) should be done first

16