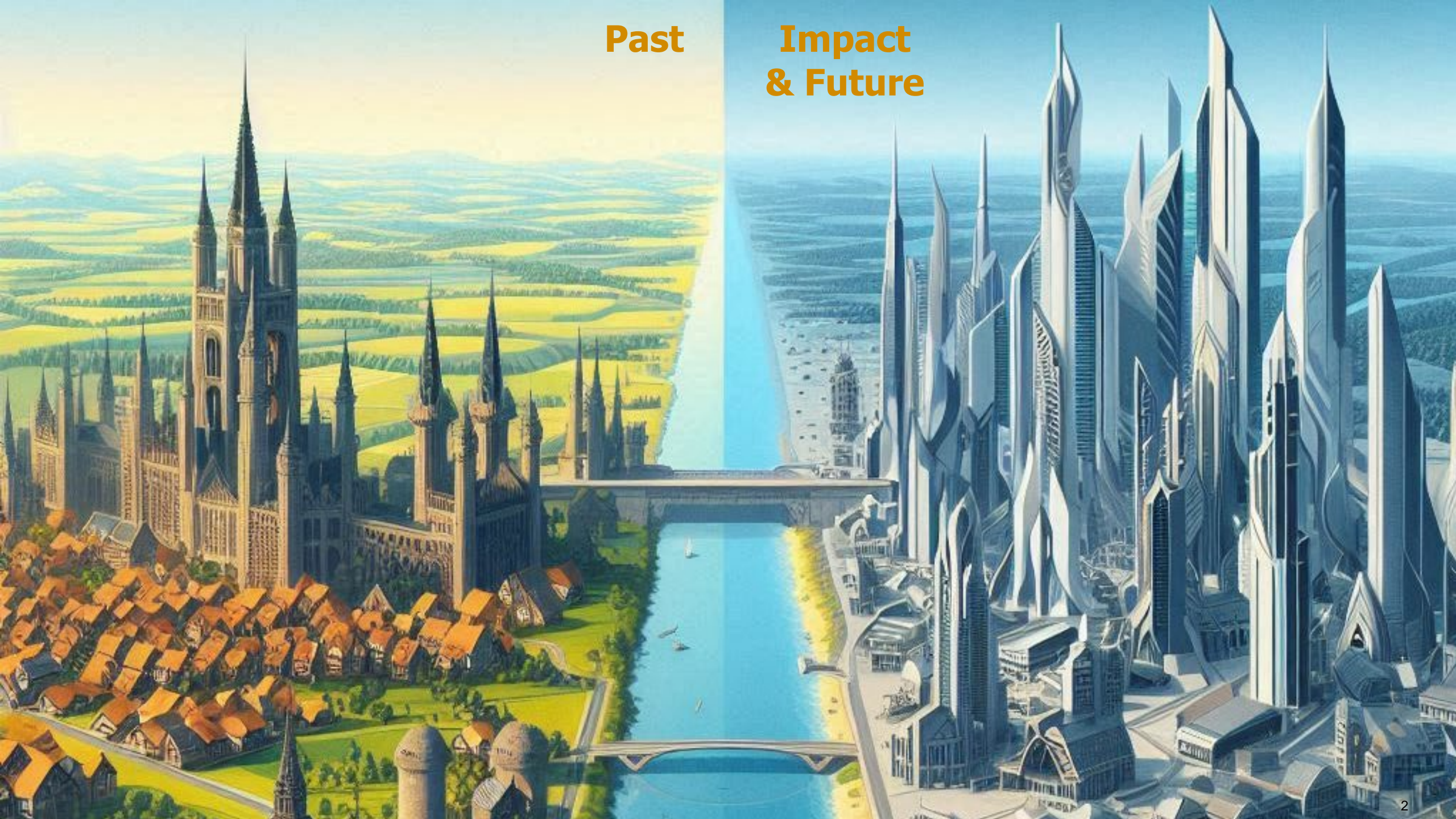
Deep Learning for Just-in-Time Defect Prediction

Most Influential Paper Award Talk

Xinli Yang, David Lo, Xin Xia, Yun Zhang, and Jianling Sun

QRS 2025, Hangzhou, China, July 2025


Past
Impact
& Future

Impact
& Future

Going Back a Decade

Exchange Program

Going Back a Decade

Deep Learning for Just-In-Time Defect Prediction

Xinli Yang*, David Lo†, Xin Xia*‡, Yun Zhang*, and Jianling Sun*
*College of Computer Science and Technology, Zhejiang University, Hangzhou, China
†School of Information Systems, Singapore Management University, Singapore
zdyxl@zju.edu.cn, davidlo@smu.edu.sg, {xxia, yunzhang28, sunjl}@zju.edu.cn

IEEE ReliabilitySociety

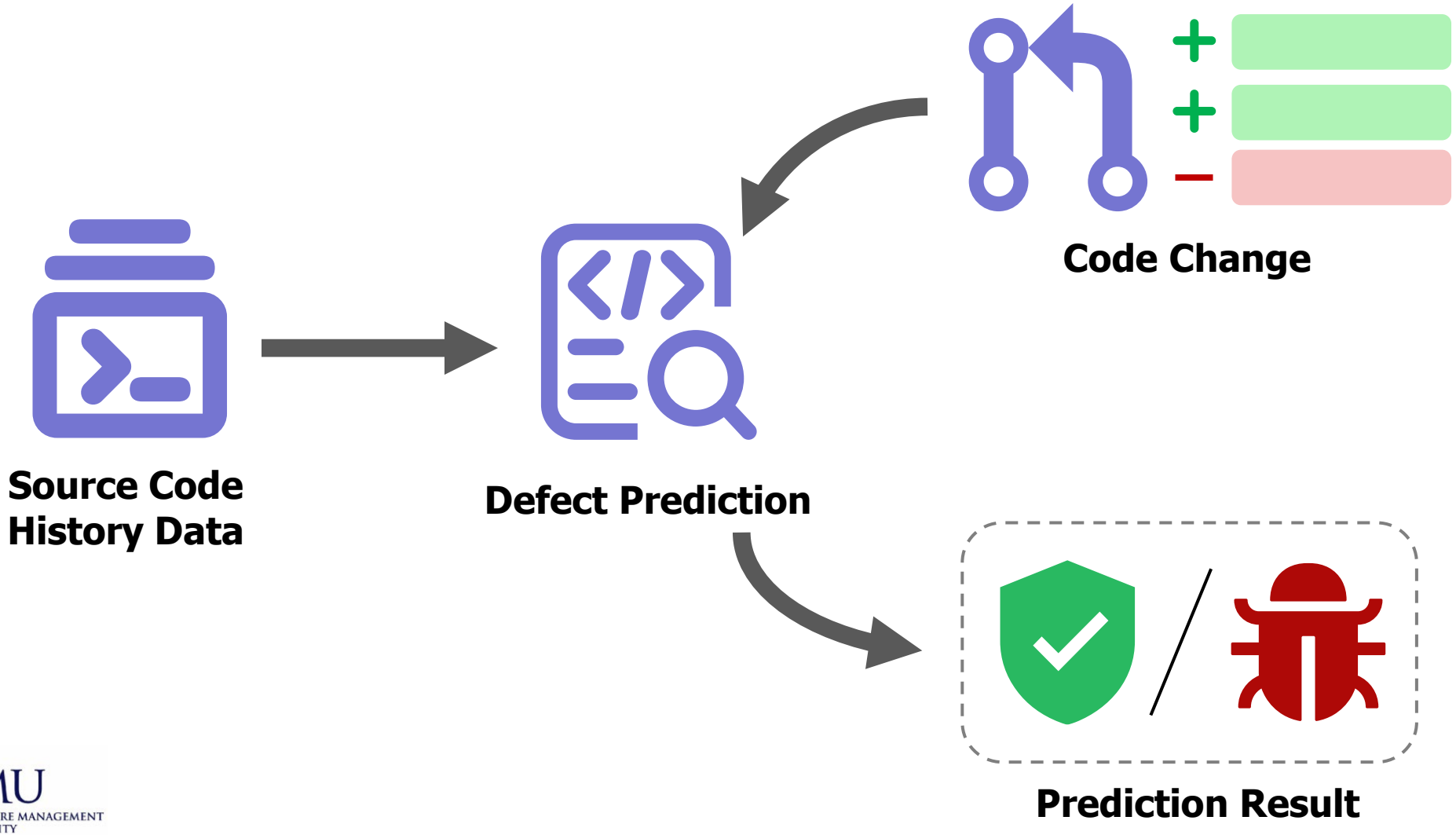
QRS 2015

The 2015 IEEE International Conference on Software Quality, Reliability & Security
Vancouver, Canada, August 3-5, 2015
http://paris.utdallas.edu/qrs15

Starting from 2015, the SERE conference (IEEE International Conference on Software Security and Reliability) and the QSIC conference (IEEE International Conference on Quality Software) will merge into one large conference QRS, with Q representing Quality, R for Reliability, and S for Security. This enhanced platform,

Just-in-time Defect Prediction

- **Predict whether a change is buggy**

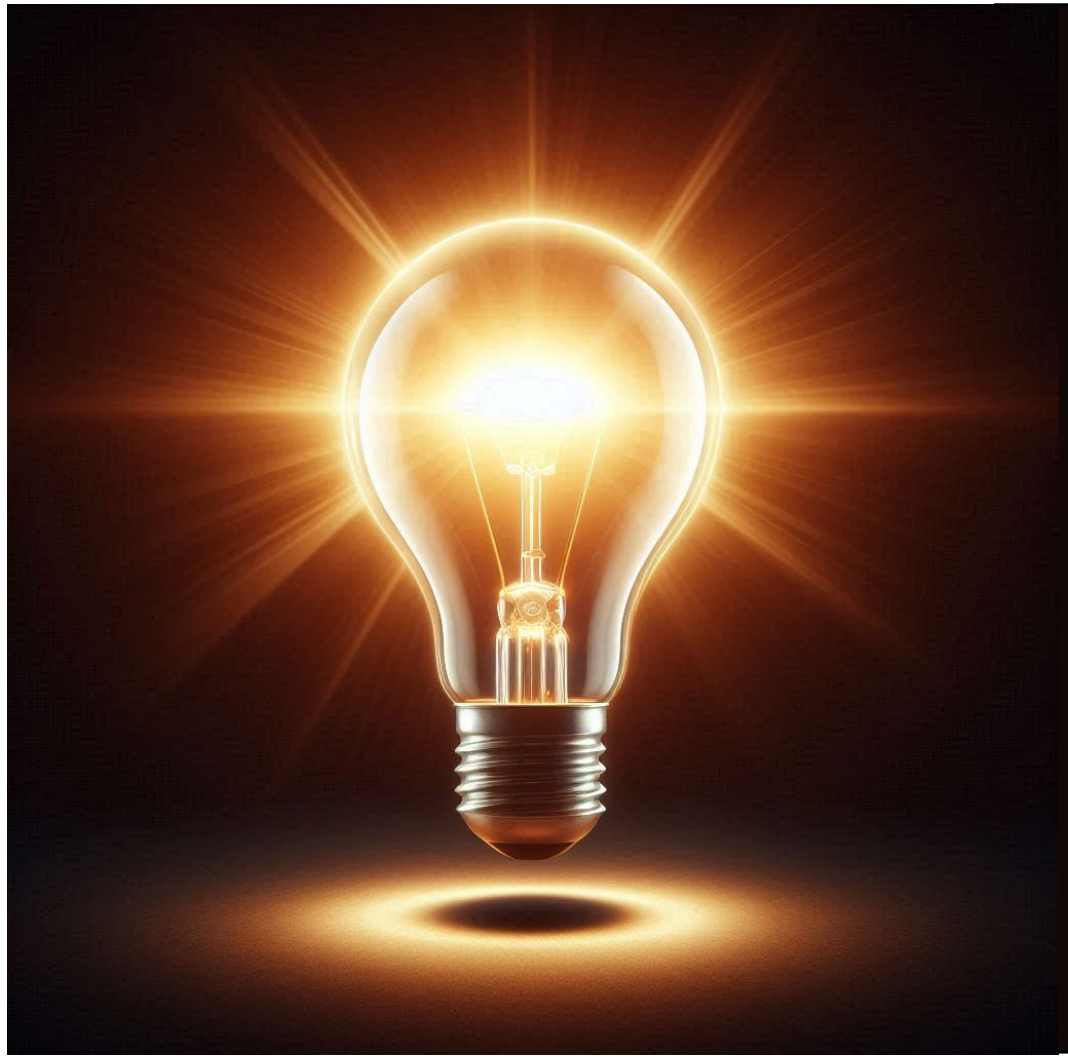


Traditional Solutions for JIT DP

- We extract a number of metrics from the historical changes with known defective information
- We build a prediction model (e.g., Random Forest) on these metrics

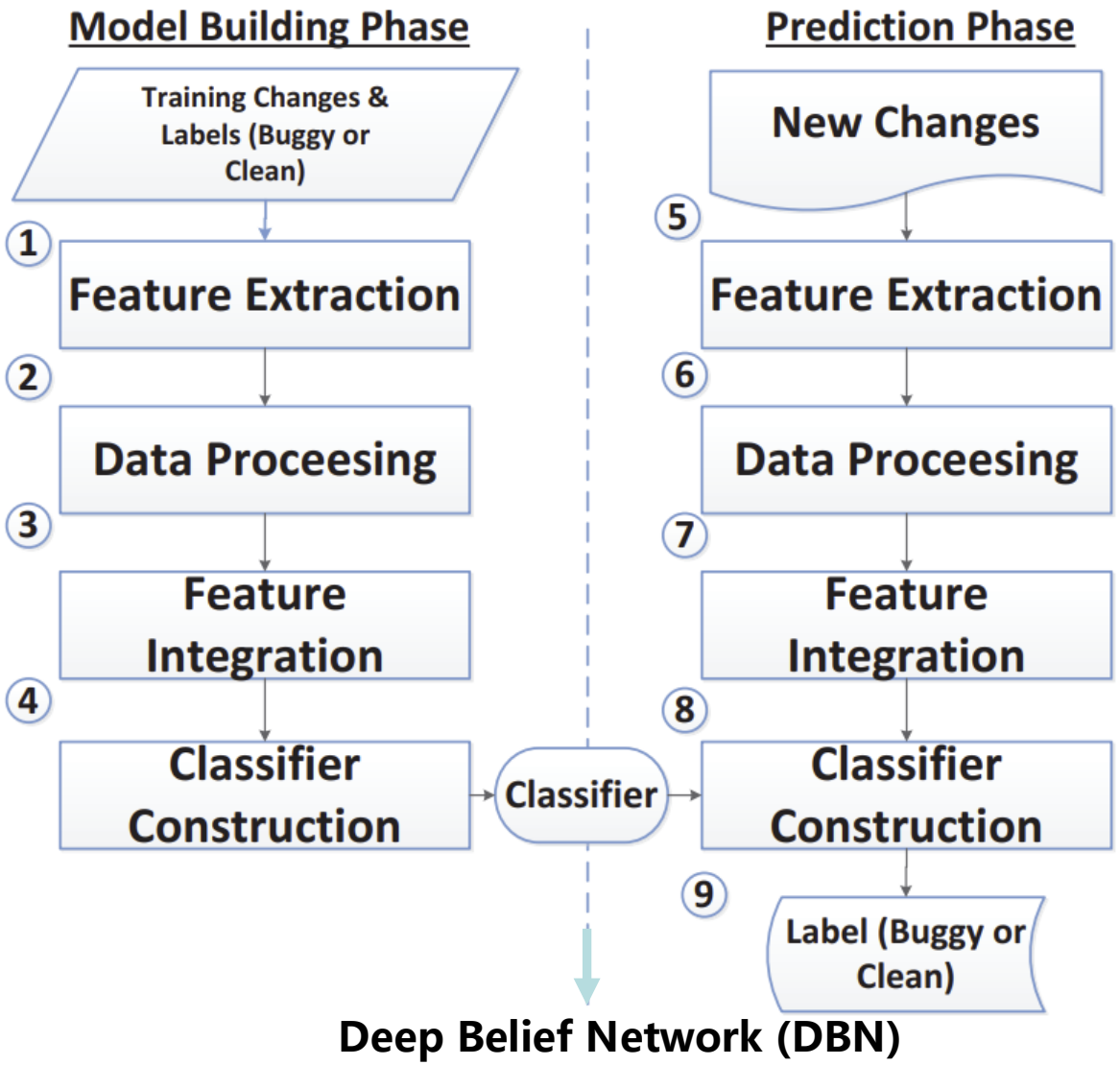
Metric	Description
NS	Number of subsystems touched by the current change
ND	Number of directories touched by the current change
NF	Number of files touched by the current change
Entropy	Distribution across the touched files
LA	Lines of code added by the current change
LD	Lines of code deleted by the current change
LT	Lines of code in a file before the current change
FIX	Whether or not the current change is a defect fix
NDEV	Number of developers that changed the files
AGE	Average time interval between the last and current change
NUC	Number of unique last changes to the files

Motivation



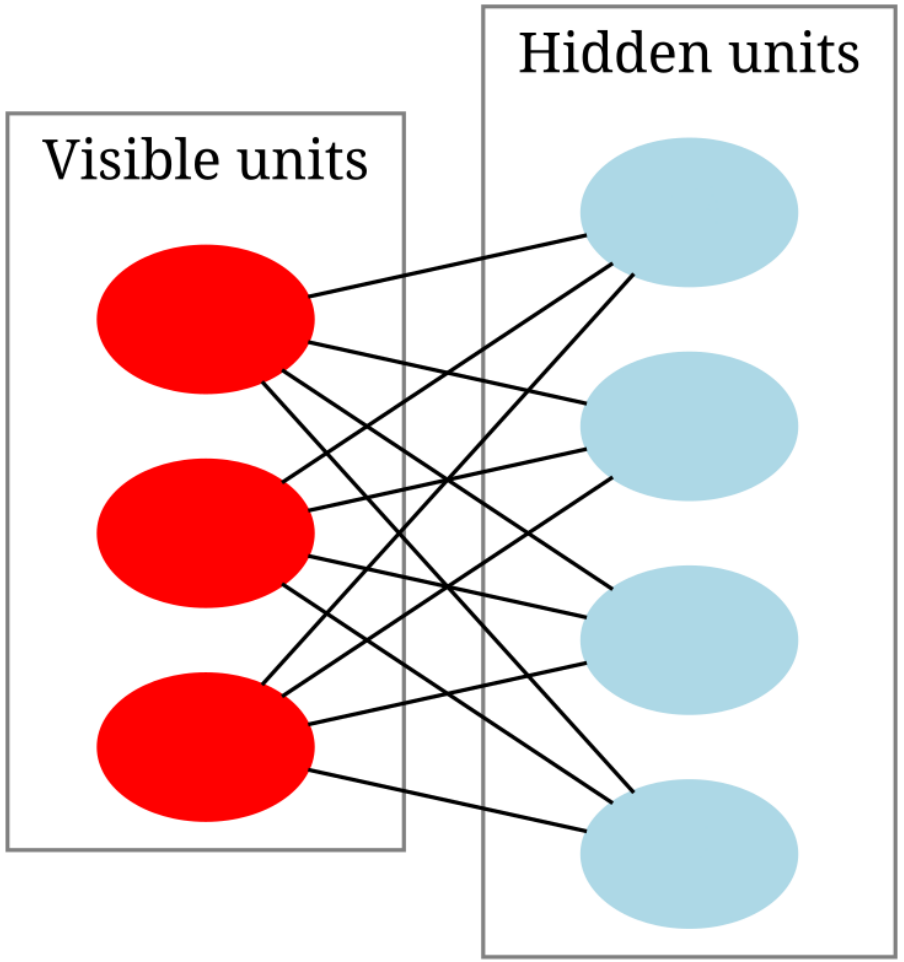
- Deep Learning has shown promising results in other area, **how about SE?**
- Features are based on expert knowledge; Can we generate advanced/hidden features based on the existing features?

Deeper: Proposed Approach



Deep Belief Network (DBN)

Structure of DBN



Experiment Design

- Six Datasets
- Two Evaluation Metrics :
  - F1-score
  - Cost-effectiveness

Project	Time	# Instances	% Buggy
Bugzilla	1998.08-2006.12	4620	36%
Columba	2002.11-2006.07	4455	31%
JDT	2001.05-2007.12	35386	14%
Platform	2001.05-2007.12	64250	14%
Mozilla	2000.01-2006.12	98275	5%
PostgreSQL	1996.07-2010.05	20431	25%

How Effective is Deeper？

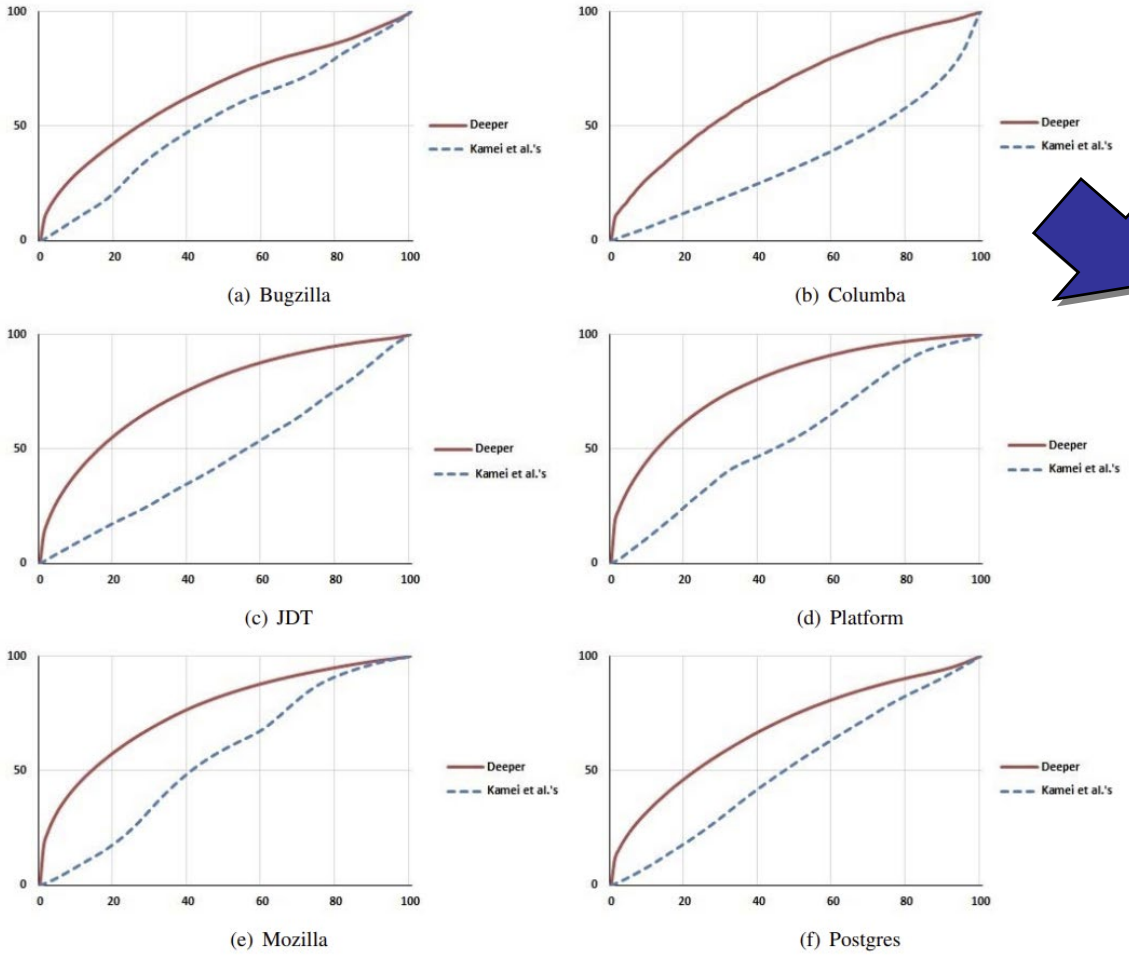
F1-scores

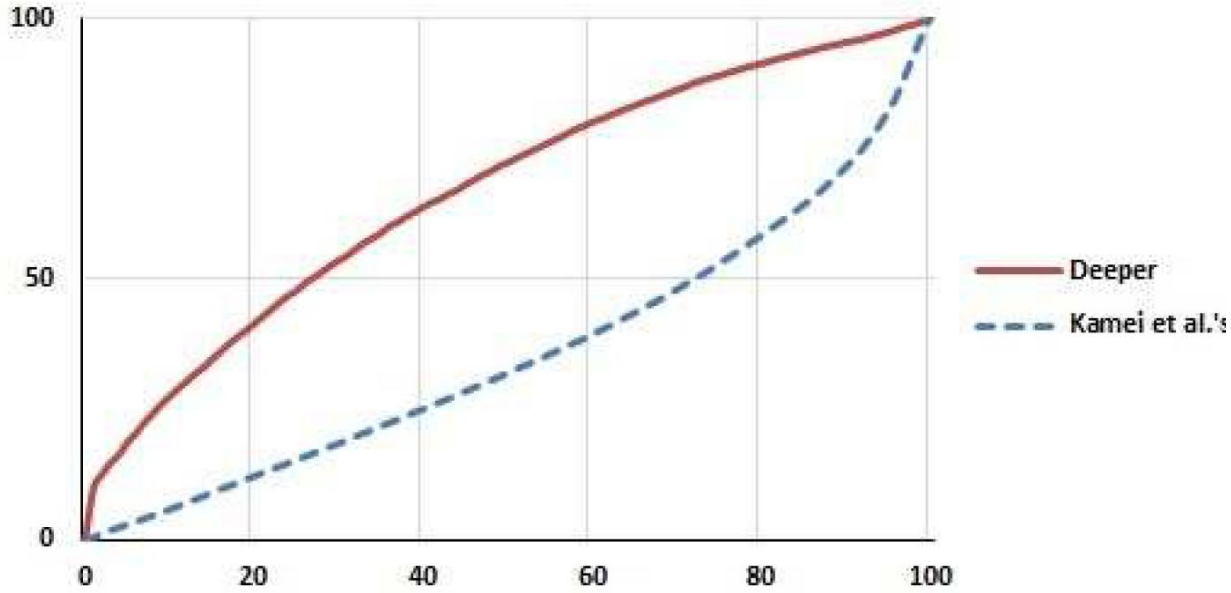
Project	LR	Kamei et al.'s	Deeper
Bugzilla	0.5106	0.6147	0.6264
Columba	0.4148	0.5550	0.5493
JDT	0.0568	0.3616	0.3769
Platform	0.0603	0.3496	0.3833
Mozilla	0.0742	0.2058	0.2213
PostgreSQL	0.4014	0.5480	0.5463
Average	0.2530	0.4391	0.4506

Cost-effectiveness

Project	LR(%)	Kamei et al.'s (%)	Deeper(%)
Bugzilla	21.35	21.44	42.80
Columba	12.52	12.29	41.00
JDT	18.31	17.79	55.77
Platform	25.71	24.92	61.87
Mozilla	18.85	18.13	58.09
PostgreSQL	17.82	18.38	46.70
Average	19.09	18.82	51.04

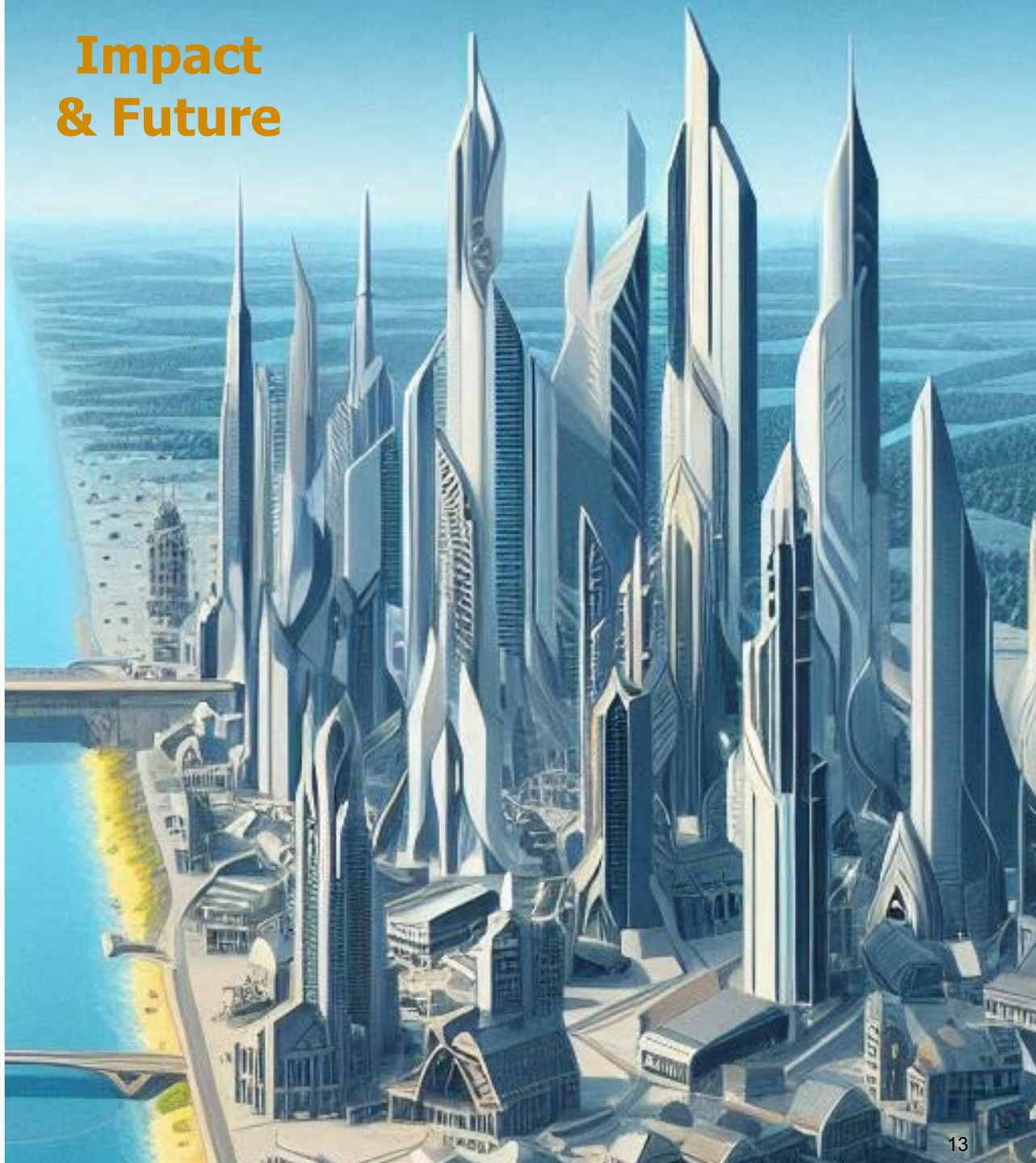
How Effective is Deeper when Different % of LOC Are Inspected?





Past

Impact & Future

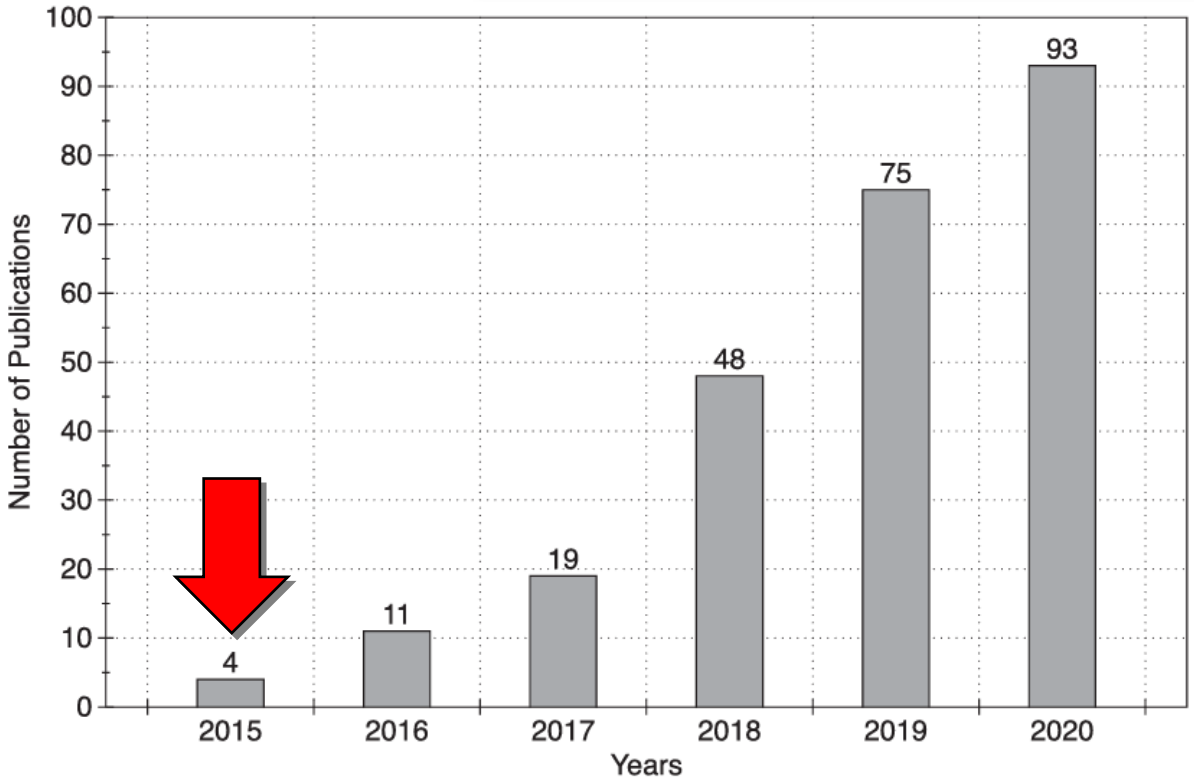


Early Work on Deep Learning for Software Engineering (DL4SE)

ACM Computing Surveys 2022

A Survey on Deep Learning for Software Engineering

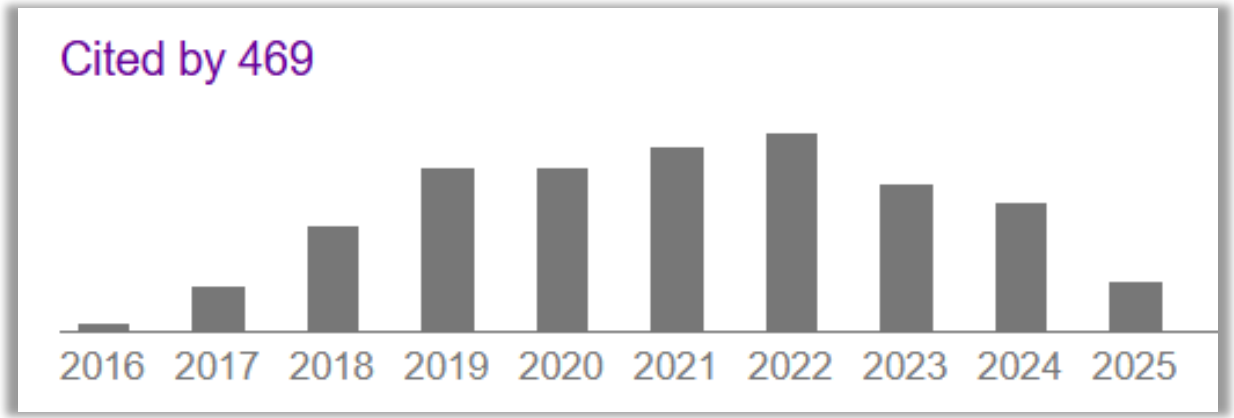
YANMING YANG, School of Computer Science and Technology, Zhejiang University, China
XIN XIA, Software Engineering Application Technology Lab, Huawei, China
DAVID LO, School of Information Systems, Singapore Management University, Singapore
JOHN GRUNDY, Faculty of Information Technology, Monash University, Australia



1. Piech et al. Learning program embeddings to propagate feedback on student code. ICML 2015 (253 citations)
2. López-Martín and Alain Abran. Neural networks for predicting the duration of new software projects. JSS 2015 (72 citations)
3. White et al. Toward deep learning software repositories. MSR 2015 (410 citations, MSR MIP)
4. Yang et al. Deep learning for just-in-time defect prediction. QRS 2015 (469 citations, QRS MIP)

First paper on deep learning for software quality, reliability, and security

Influence of the Work



I. Defect Prediction
II. Vulnerability Detection
III. Representation Learning of Software Artifacts (SA)
IV. Surveys in the Areas of AI, SE, and Cybersecurity
V. Others (code review, bug localization, build outcome prediction, technical debt prediction, effort estimation, migration, program repair, etc.)

Selected Follow Up Works – Defect Prediction

Automatically Learning Semantic Features for Defect Prediction

Song Wang,Taiyue Liu and Lin Tan
Electrical and Computer Engineering, University of Waterloo, Canada
{song.wang, t67liu, lintan}@uwaterloo.ca

ICSE 2016
900+ Citations

"**Yang et al. [68]** proposed an approach that leveraged deep learning to generate features from existing features and then used these new features to predict whether a commit is buggy or not. ... Our work differs from the above study mainly in three aspects. **First**, we use DBN to learn semantic features directly from source code ... **Second**, we evaluate the effectiveness of our generated features using different classifiers and for both within-project and cross-project defect prediction ... **Third**, we focus on file level defect prediction, while they work on change level defect prediction."

Selected Follow Up Works – Vulnerability Detection

VulDeePecker: A Deep Learning-Based System for Vulnerability Detection

Zhen Li*†, Deqing Zou*‡♯, Shouhuai Xu§, Xinyu Ou*, Hai Jin*,
Sujuan Wang*, Zhijun Deng* and Yuyi Zhong*
*Services Computing Technology and System Lab, Big Data Technology and System Lab,
Cluster and Grid Computing Lab, School of Computer Science and Technology,
Huazhong University of Science and Technology
deqingzou@hust.edu.cn
†School of Cyber Security and Computer, Hebei University
‡Shenzhen Huazhong University of Science and Technology Research Institute
§Department of Computer Science, University of Texas at San Antonio

NDSS 2018
1000+ citations

"Somewhat related work is the **use of deep learning for software defect prediction** [54], [63]. However, software defects are different from software **vulnerabilities** … Moreover, the defect prediction method presented in [63] is geared towards code changes rather than **target programs as a whole**."

[63] = QRS'15 paper, [54] = ICSE'16 paper

Selected Follow Up Works – Representation Learning of SA

ICSE 2020
250+ citations

CC2Vec: Distributed Representations of Code Changes

Thong Hoang, Hong Jin Kang, David Lo
Singapore Management University, Singapore
{vdthoang.2016,hjkang.2018,davidlo}@smu.edu.sg

Julia Lawall
Sorbonne University/Inria/LIP6, France
Julia.Lawall@inria.fr

"These models often **do not model the hierarchical structure** of a code change or require handcrafted features that may be **specific to a single task** [... 60]."

[60] = Our QRS'15 paper

Future Directions

Foundation Model-Based Agents

Let the Trial Begin: A Mock-Court Approach to Vulnerability Detection using LLM-Based Agents

Ratnadira Widyasari
ratnadiraw@smu.edu.sg
Singapore Management University
Singapore

Martin Weyssow
mweyssow@smu.edu.sg
Singapore Management University
Singapore

Ivana Clairine Irsan
ivanairsan@smu.edu.sg
Singapore Management University
Singapore

Han Wei Ang
ang_han_wei@tech.gov.sg
GovTech
Singapore

Frank Liauw
frank_liauw@tech.gov.sg
GovTech
Singapore

Eng Lieh Ouh
elouh@smu.edu.sg
Singapore Management University
Singapore

Lwin Khin Shar
lkshar@smu.edu.sg
Singapore Management University
Singapore

Hong Jin Kang
hongjin.kang@sydney.edu.au
University of Sydney
Australia

David Lo
davidlo@smu.edu.sg
Singapore Management University
Singapore

2025 FORGE

The 2nd ACM International Conference on AI Foundation Models and Software Engineering

Sun 27 - Mon 28 April 2025 Ottawa, Ontario, Canada

Future Directions

Human-Centered AI

Explaining Explanations: An Empirical Study of Explanations in Code Reviews

RATNADIRA WIDYASARI and TING ZHANG, School of Computing and Information Systems, Singapore Management University, Singapore, Singapore
ABIR BOURAFFA and WALID MAALEJ, University of Hamburg, Hamburg, Germany
DAVID LO, School of Computing and Information Systems, Singapore Management University, Singapore, Singapore

TOSEM 2025

The 1st Workshop on Human-Centered AI for SE

"Where AI4SE Meets Human Insight"

HumanAISE Workshop (Co-located with FSE'25 and ISSTA'25 at Trondheim, Norway).

Thank You!

- QRS Steering Committee
- QRS'25 MIP Award Committee
- QRS'15 PC Chairs and Members
- Everyone who have inspired us, collaborated with us, and extended our work

SMU SINGAPORE MANAGEMENT UNIVERSITY
OUB Chair Professorship Fund

Going Back a Decade

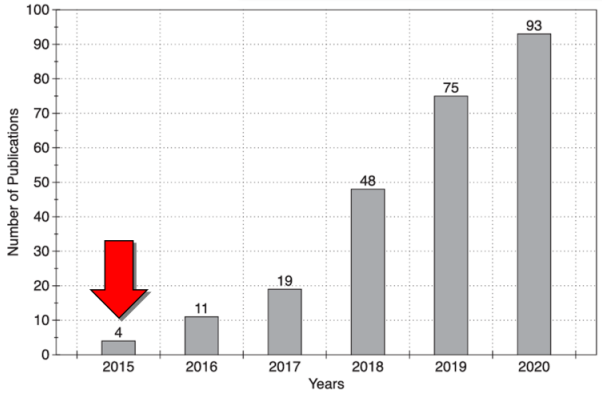
Exchange Program

Early Work on Deep Learning for Software Engineering (DL4SE)

ACM Computing Surveys 2022

A Survey on Deep Learning for Software Engineering

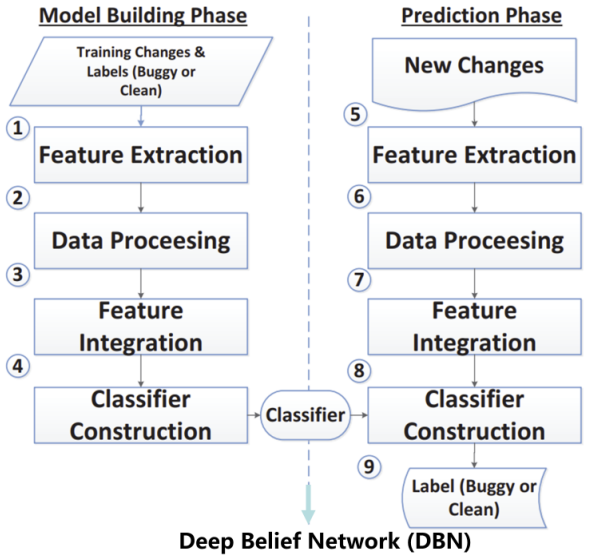
YANMING YANG, School of Computer Science and Technology, Zhejiang University, China
XIN XIA, Software Engineering Application Technology Lab, Huawei, China
DAVID LO, School of Information Systems, Singapore Management University, Singapore
JOHN GRUNDY, Faculty of Information Technology, Monash University, Australia



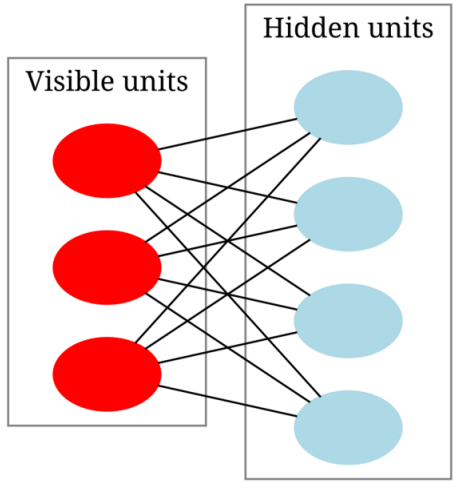
1. Piech et al. Learning program embeddings to propagate feedback on student code. ICML 2015 (253 citations)
2. López-Martín and Alain Abran. Neural networks for predicting the duration of new software projects. JSS 2015 (72 citations)
3. White et al. Toward deep learning software repositories. MSR 2015 (410 citations, MSR MIP)
4. Yang et al. Deep learning for just-in-time defect prediction. QRS 2015 (469 citations, QRS MIP)

First paper on deep learning for software quality, reliability, and security

Deeper: Proposed Approach



Structure of DBN



Future Directions

Foundation Model-Based Agents

2025 FORGE

The 2nd ACM International Conference on AI Foundation Models and Software Engineering

Sun 27 - Mon 28 April 2025 Ottawa, Ontario, Canada

Human-Centered AI

The 1st Workshop on Human-Centered AI for SE

"Where AI4SE Meets Human Insight"

HumanAISE Workshop (Co-located with FSE'25 and ISSTA'25 at Trondheim, Norway).

Thank You!

Questions? Comments? Advice?

xin.xia@acm.org

davidlo@smu.edu.sg

Going Back a Decade

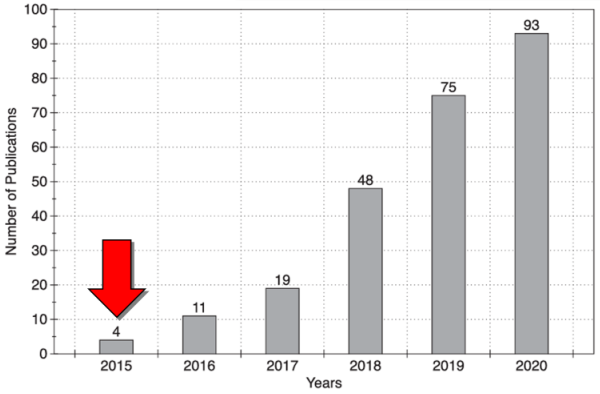
Exchange Program

Early Work on Deep Learning for Software Engineering (DL4SE)

ACM Computing Surveys 2022

A Survey on Deep Learning for Software Engineering

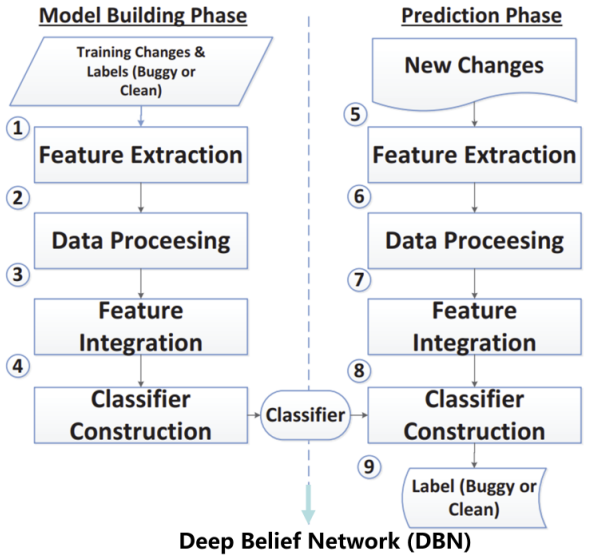
YANMING YANG, School of Computer Science and Technology, Zhejiang University, China
XIN XIA, Software Engineering Application Technology Lab, Huawei, China
DAVID LO, School of Information Systems, Singapore Management University, Singapore
JOHN GRUNDY, Faculty of Information Technology, Monash University, Australia



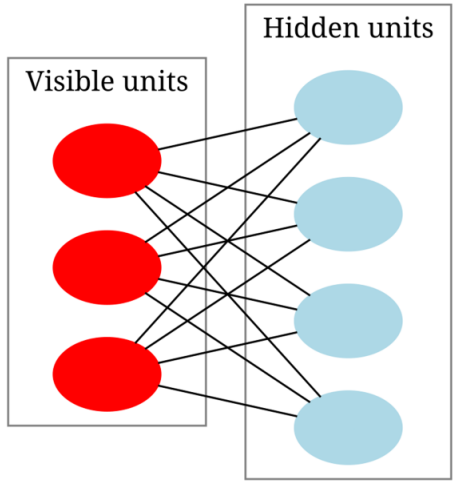
1. Piech et al. Learning program embeddings to propagate feedback on student code. ICML 2015 (253 citations)
2. López-Martín and Alain Abran. Neural networks for predicting the duration of new software projects. JSS 2015 (72 citations)
3. White et al. Toward deep learning software repositories. MSR 2015 (410 citations, MSR MIP)
4. Yang et al. Deep learning for just-in-time defect prediction. QRS 2015 (469 citations, QRS MIP)

First paper on deep learning for software quality, reliability, and security

Deeper: Proposed Approach



Structure of DBN



Future Directions

Foundation Model-Based Agents

2025 FORGE

The 2nd ACM International Conference on AI Foundation Models and Software Engineering

Sun 27 - Mon 28 April 2025 Ottawa, Ontario, Canada

Human-Centered AI

The 1st Workshop on Human-Centered AI for SE

"Where AI4SE Meets Human Insight"

HumanAISE Workshop (Co-located with FSE'25 and ISSTA'25 at Trondheim, Norway).